



**UNIVERSIDADE FEDERAL DE SERGIPE
CENTRO DE CIÊNCIAS SOCIAIS E APLICADAS
NÚCLEO DE CIÊNCIA DA INFORMAÇÃO
CURSO DE BIBLIOTECONOMIA E
DOCUMENTAÇÃO**



VANESSA REIS BARROS

**TECNOLOGIAS PARA A REPRESENTAÇÃO DESCRITIVA NA *WEB* SEMÂNTICA
E SUAS CONVERGÊNCIAS COM AS BIBLIOTECAS DIGITAIS**

SÃO CRISTOVÃO

2013

VANESSA REIS BARROS

**TECNOLOGIAS PARA A REPRESENTAÇÃO DESCRITIVA NA *WEB* SEMÂNTICA
E SUAS CONVERGÊNCIAS COM AS BIBLIOTECAS DIGITAIS**

Trabalho de Conclusão de Curso apresentado
ao Núcleo de Ciência da Informação, da
Universidade Federal de Sergipe, para
obtenção do grau de bacharel em
Biblioteconomia e Documentação.

Orientador: Prof. Dr. Fabiano Ferreira de
Castro.

Linha de Pesquisa: Informação e Tecnologia.

SÃO CRISTOVÃO

2013

Barros, Vanessa Reis.

B2786t Tecnologias para a representação descritiva na web
semântica e suas convergências com as bibliotecas digitais /
Vanessa Reis Barros. - São Cristóvão, 2013.

88 f. : il.

Trabalho de Conclusão de Curso (Graduação em
Biblioteconomia e Documentação) – Núcleo de Ciência da
Informação – Universidade Federal de Sergipe, 2013.

Orientador: Dr. Fabiano Ferreira de Castro.

1. Web Semântica. 2. Bibliotecas Digitais. 3. Camadas – Web
Semântica 4. Padrões de Metadados. I. Autor. II. Título.

CDU 02 (004.62)

TECNOLOGIAS PARA A REPRESENTAÇÃO DESCRITIVA NA *WEB SEMANTICA* E SUAS CONVERGÊNCIAS COM AS BIBLIOTECAS DIGITAIS

VANESSA REIS BARROS

Trabalho de Conclusão de Curso apresentado
ao Núcleo de Ciência da Informação, da
Universidade Federal de Sergipe, para
obtenção do grau de bacharel em
Biblioteconomia e Documentação.

Nota: _____

Data de Apresentação: _____

BANCA EXAMINADORA

Prof. Dr. Fabiano Ferreira de Castro
(Orientador)

Prof. Dr. Jefferson David Araújo Sales
(Membro Convidado - Externo)

Prof. Dra. Valéria Aparecida Bari
(Membro Convidado - Interno)

AGRADECIMENTOS

Agradeço primeiramente a Deus, por nos dar o dom da vida e o ânimo para desempenhar atividades cotidianas.

Dedico este espaço também para agradecer a minha família, principalmente a minha mãe, dona Lúcia, que em muito ajudou nos momentos de dificuldades, dúvidas e também de alegrias em todo o decorrer do curso e da vida.

Ao corpo docente do Núcleo de Ciência da Informação da UFS que aceitou o desafio de apresentar o “mundo da biblioteconomia” para os alunos da primeira turma formada pela Universidade Federal de Sergipe, contribuindo assim para nossa formação.

À Profa. Dra. Valéria Aparecida Bari, que contribuiu decisivamente para o desenvolvimento do curso.

Ao meu orientador, o Professor Doutor Fabiano Ferreira de Castro, que muito me auxiliou na composição deste trabalho.

Além de todos os amigos, colegas e pessoas sempre muito especiais que estiveram comigo neste período, quero dizer um muito obrigada a todos.

“Que os vossos esforços desafiem as impossibilidades, lembrai-vos de que as grandes coisas do homem foram conquistadas do que parecia impossível.”

Charles Chaplin

RESUMO

O surgimento da *Web* foi um marco para a disseminação de conhecimento, a cada dia novos usuários passam a ter acesso a uma gama de conteúdos que seria impossível de encontrar em um único ambiente informacional. No cenário atual é possível identificar o desenvolvimento de conteúdos em formatos digitais e a necessidade de recuperá-los de forma eficiente é uma constante. Algumas iniciativas são criadas para o estabelecimento de tecnologias que propiciam a representação e a descrição dos dados, no intuito de potencializar a recuperação. As bibliotecas digitais que são consideradas ambientes colaborativos são apontadas como uma alternativa para a construção da *Web Semântica*. Objetiva-se estudar as camadas iniciais da estrutura da *Web Semântica*, onde estão concentradas formas de representação descritiva, linguagens e padrões de metadados para avaliar as semelhanças e a possibilidade de implantação ou agregação às já existentes no desenvolvimento de bibliotecas digitais. A metodologia caracteriza-se por ser de análise exploratória e descritiva do tema, utilizando-se o método de pesquisa bibliográfica e documental, estudando os principais conceitos relacionados ao tema de pesquisa, tais como Bibliotecas digitais, *Web Semântica*, Representação descritiva e Metadados, na literatura científica das áreas de Ciência da Informação, Biblioteconomia e Ciência da Computação, a fim de verificar as tecnologias computacionais e informacionais que fundamentam as bibliotecas digitais na construção *Web Semântica*. Assim, é possível identificar que há uma sinergia entre bibliotecas digitais e *Web Semântica*, pois a primeira contribui com seu ambiente padronizado, pautado nas metodologias da Biblioteconomia, tais como a representação descritiva; a segunda oferece nas tecnologias computacionais semânticas, a possibilidade de potencializar os serviços oferecidos pelas bibliotecas digitais, potencializando a busca e agregando maior semântica aos recursos informacionais, facilitando a recuperação pelos agentes computacionais e apresentando informações consistentes e significativas para o usuário final.

Palavras-chave: *Web Semântica*. Bibliotecas Digitais. Representação Descritiva. Metadados. Informação e Tecnologia.

ABSTRACT

The emergence of the Web has been a landmark for the dissemination of knowledge, every day new users will have access to a range of content that would be impossible to find in a single information environment. In the current scenario it is possible to identify the development of content in digital formats and the need to retrieve them efficiently is a constant. Some initiatives are created for the development of technologies that provide the representation and description of the data in order to maximize recovery. Digital libraries are considered collaborative environments are seen as an alternative to building the Semantic Web. The objective is to study the structure of the initial layers of the Semantic Web, where they are concentrated forms of descriptive representation, languages and metadata standards to assess the similarities and the possibility of deployment to existing or aggregation in the development of digital libraries. The methodology is characterized by being exploratory and descriptive analysis of the topic, using the method of documentary and bibliographical research, studying the main concepts related to the research topic, such as digital libraries, Semantic Web, Metadata and Descriptive Representation in scientific literature in the areas of Information Science, Library and Computer Science in order to verify the computational and informational technologies underlying digital libraries in building the Semantic Web. Thus, it is possible to identify that there is a synergy between digital libraries and the Semantic Web, as the first contributes its standardized environment, based on the methodologies of librarianship, such as descriptive representation and the second semantic computing technologies offers the possibility to enhance the services offered by digital libraries, leveraging searching and adding more semantics to information resources, facilitating recovery by computational agents and presenting accurate and meaningful information to the end user.

Keywords: Semantic Web. Digital Libraries. Descriptive representation. Metadata. Information and Technology.

LISTA DE GRÁFICOS

Gráfico 1 - Comparativo percentual das pessoas que utilizaram a Internet nos últimos 3 meses, brasileiros, faixa etária acima dos 9 anos, segundo escolaridade – 2005/2008	20
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----

LISTA DE FIGURAS

Figura 1 – Página Inicial da Biblioteca Digital Mundial	32
Figura 2 – Página Inicial da Biblioteca Digital Brasileira de Teses e Dissertações – BDTD	32
Figura 3 – Comparativo entre serviços da web 2.0 e serviços e serviços da biblioteca tradicional	34
Figura 4 - Características comuns simultaneamente à <i>Web</i> e Bibliotecas 1.0 e 2.0.....	37
Figura 5 – Evolução da Web 1.0 a 4.0	53
Figura 6 - Arquitetura em camadas da <i>Web</i> Semântica	56
Figura 7 – Códigos da Tabela ASCII	60
Figura 8 – Comparativo entre os Formatos de Transformação Unicode	61
Figura 9 – Código em Linguagem HTML	64
Figura 10 – Exemplo de código XML.....	64
Figura 11 – Exemplo de código em RDF	66
Figura 12 – Exemplo de MARC 21	70

LISTA DE QUADROS

Quadro 1 - Principais tecnologias que representam a evolução da <i>Web</i>	28
Quadro 2 - Evolução da Biblioteca 1.0 para Biblioteca 2.0	30
Quadro 3 – Habilidades profissionais necessárias em Bibliotecas Digitais.....	49
Quadro 4 – Elementos básicos do padrão Dublin Core	72

LISTA DE SIGLAS

CSS – Cascading Style Sheets

HTML – HiperText Markup Language

RDF - Resource Description Framework

RSS – Really Simple Syndication

SGML - Standard Generalized Markup Language

TIC – Tecnologia da Informação e Comunicação

URI - Uniform Resource Identifier

URL – Uniform Resource Locator

W3C - World Wide Web Consortium

WWW - World Wide Web

XML – eXtensible Markup Language

SUMÁRIO

1	INTRODUÇÃO.....	14
1.1.	Motivação.....	14
1.2.	Objetivos.....	15
1.2.1.	Objetivo Geral.....	15
1.2.2.	Objetivos Específicos	15
1.3.	Justificativa	16
2	TRANSFORMAÇÕES DA <i>WEB</i> E SUA RELAÇÃO COM OS BIBLIOTECÁRIOS.....	18
2.1	A evolução da <i>Web</i>	21
2.1	Características atribuídas as <i>Web</i> 1.0 e <i>Web</i> 2.0.....	24
2.2	A Biblioteconomia no contexto <i>Web</i>	28
2.3	A catalogação e suas contribuições para a <i>Web</i>	35
3	BIBLIOTECAS DIGITAIS: contexto, conceitos e tecnologias	38
3.1	A Biblioteca Universal	40
3.2	Biblioteca virtual ou digital?	43
3.3	Afinal, o que é Biblioteca Digital?	44
3.3.1	Conceitos atribuídos à Biblioteca Digital	44
3.3.2	Metodologia para a construção de Bibliotecas Digitais	47
4	AS TECNOLOGIAS DAS CAMADAS-BASE DA <i>WEB</i> SEMÂNTICA E SUAS CORRELAÇÕES COM AS BIBLIOTECAS DIGITAIS	51
4.1	Acerca da <i>Web</i> Semântica.....	51
4.2	Estrutura da <i>Web</i> Semântica.....	56
4.2.1	Camada Internacional (Unicode/ URI)	58
4.2.1.1	O Padrão Unicode	58
4.2.1.2	Uniform Resource Identifiers – URI	62
4.2.2	Camada Sintática (<i>XML</i> e <i>Namespace</i>)	63
4.2.3	Camada de dados (<i>RDF</i> e <i>RDF Schema</i>).....	66
4.3	O que as Bibliotecas Digitais têm em comum com as camadas-base da <i>Web</i> Semântica?.....	69
4.3.1	O Formato MARC 21	69
4.3.2	O padrão de metadados Dublin Core.....	72
5	CONSIDERAÇÕES FINAIS	Erro! Indicador não definido.
	REFERÊNCIAS	81

1 INTRODUÇÃO

Esta pesquisa está inserida na linha de pesquisa Informação e Tecnologia do Núcleo de Ciência da Informação (NUCI), da Universidade Federal de Sergipe e vinculada ao Grupo de Estudos e Pesquisas em Representação do Conhecimento e Tecnologias da Informação e Comunicação (GPERTIC).

O objeto de investigação consiste no estudo das semelhanças e possíveis aplicações dos instrumentos tecnológicos utilizados nas camadas iniciais da arquitetura da *Web Semântica*, para implantá-los no desenvolvimento das bibliotecas digitais, visando com isto, aperfeiçoar os serviços proporcionados por este tipo de biblioteca e mostrar a proximidade na forma de organização sugerida para esta *Web*.

Para tanto, surge o problema de pesquisa: As bibliotecas digitais podem contribuir para o desenvolvimento da *Web Semântica*? Poderia a *Web Semântica* contribuir no estabelecimento de bibliotecas digitais?

O presente trabalho caracteriza-se por ser uma pesquisa de análise exploratória e descritiva do tema, utilizando-se o método de pesquisa bibliográfica e documental, estudando os principais conceitos relacionados ao tema de pesquisa, tais como Bibliotecas digitais, *Web Semântica*, Representação descritiva e Metadados, na literatura científica das áreas de Ciência da Informação, Biblioteconomia e Ciência da Computação, a fim de verificar as tecnologias computacionais e informacionais que fundamentam as bibliotecas digitais na construção *Web Semântica*.

É dada prioridade às produções científicas entre o período de 2001 a 2013, observando que a data inicial é coincidente com a publicação do artigo de Berners-Lee, Hendler e Lassila (2001) sobre a *Web Semântica*, texto este, que divulgou esta nova estrutura de modo mais efetivo para os cientistas de todo o mundo.

1.1. Motivação

Esta pesquisa foi motivada pelo que se observa na literatura científica, a resistência de profissionais da área da Biblioteconomia em associar o trabalho cotidiano do bibliotecário num ambiente real (físico) com o que há de inovador nas

tecnologias advindas do ambiente virtual. Neste ponto, uma curiosidade estudantil causou a escolha do tema Biblioteca Digital e *Web Semântica*.

Foram escolhidas as três camadas iniciais da *Web Semântica*, em virtude destas estarem mais associadas à representação e à descrição de dados, para demonstrar que, o que se propõe como novidade no meio digital já seria utilizado nas bibliotecas, com outros termos, relativos à catalogação, indexação, organização, etc.

Quando se pensa em bibliotecas na internet, padrões e metadados, fica ainda mais fácil de visualizar uma proximidade tecnológica entre elas, o que torna interessante um estudo das práticas recomendadas para o desenvolvimento de ambientes informacionais digitais semânticos.

Além disso, observar se tais camadas realmente convergem com o que é utilizado nas unidades de informação digitais, pensando desta forma, veremos o que cada uma pode contribuir para a melhoria de desempenho da outra, visando o mesmo objetivo de melhorar os serviços oferecidos aos usuários da *Web*.

1.2. Objetivos

1.2.1. Objetivo Geral

Estudar os conceitos e padrões utilizados nas bibliotecas digitais, as camadas iniciais da *Web Semântica* na perspectiva da Ciência da Informação e sua aplicabilidade nas bibliotecas digitais.

1.2.2. Objetivos Específicos

- ◆ Conceituar a *Web Semântica* sob a perspectiva das áreas da Ciência da Informação e da Ciência da Computação;
- ◆ Identificar o funcionamento da *Web Semântica* a partir das camadas de representação e descrição de recursos informacionais;
- ◆ Identificar as principais tecnologias utilizadas na *Web Semântica* e nas bibliotecas digitais;
- ◆ Verificar em que medida as tecnologias utilizadas na *Web Semântica* contribuem para o desenvolvimento de bibliotecas digitais.

1.3. Justificativa

Este trabalho se justifica pelas contribuições teóricas trazidas para a linha de pesquisa, no qual está inserida Informação e Tecnologia, do Núcleo de Ciência da Informação, da Universidade Federal de Sergipe.

Além disso, a proposta de trazer um estudo de temas em sua fase emergente (ou pelo menos onde não se tem conceitos já estabilizados), foco de discussões de diversas áreas do conhecimento e difícil conceituação.

A relevância social se destaca por fornecer subsídios para o desenvolvimento de bibliotecas digitais sob os moldes semânticos, proporcionando sistemas mais eficientes; visa mostrar a colaboração mútua entre padrões da *Web Semântica* e Ciência da Informação e também servir de incentivo para elaboração de mais pesquisas nesta área, sobretudo, no estado de Sergipe.

Para tanto, este Trabalho de Conclusão de Curso está estruturado, além do capítulo 1, que aborda questões introdutórias da investigação, como tema, problemas de pesquisa, objetivos, metodologia e justificativa, da seguinte maneira:

No capítulo 2 será mostrado o contexto no qual surgiu a *Web Semântica*, através da evolução da mesma e a importância das bibliotecas durante os tempos, além da sua inserção no ambiente virtual acompanhando o desenvolvimento das Tecnologias da Informação e Comunicação (TIC). Também são apresentados os conceitos relativos à *World Wide Web*, sua transformação desde o seu surgimento passando pela *Web Sintática* e uma breve introdução à terceira geração traçando, um paralelo com a Biblioteconomia, que sempre se fez presente no meio digital.

No capítulo 3 serão debatidas questões sobre as polêmicas bibliotecas digitais, que recebem diferentes nomenclaturas e significados dependendo da linha seguida por alguns autores. Também haverá uma pequena discussão sobre o que é virtual e o antigo sonho da humanidade de armazenar toda a informação produzida em um só lugar, a chamada biblioteca universal, que atualmente está sendo comparada à própria *Web*.

No capítulo 4 serão discutidas as terminologias atribuídas à *Web Semântica*, conhecendo um pouco mais sobre sua estrutura em camadas, direcionando para os três pilares da arquitetura, estudando as tecnologias contidas e possibilitando a observação do que é empregado para a padronização e representação de dados nas bibliotecas digitais, com o objetivo de verificar se ambas as tecnologias

realmente são semelhantes e se podem ou não ser aplicadas nestes ambientes informacionais.

No capítulo 5 são apresentadas as considerações finais, destacando, as observações, reflexões e posicionamentos sobre as tecnologias computacionais e informacionais que permeiam o desenvolvimento de bibliotecas digitais e também da *Web Semântica*.

2 TRANSFORMAÇÕES DA WEB E SUA RELAÇÃO COM OS BIBLIOTECÁRIOS

Neste capítulo trataremos sobre a evolução da *World Wide Web (WWW)*, as discussões sobre as suas nomenclaturas, as transformações sofridas por esta grande rede mundial de documentos durante sua breve história para alcançar formas cada vez mais eficientes de comunicação, auxiliando na troca de informações de modo ágil e trazendo grandes mudanças nos diversos segmentos sociais.

A cada dia novas áreas percebem as vantagens do uso do ambiente *Web*, com a Biblioteconomia não poderia ser diferente, portanto, procura-se contextualizar os profissionais da informação com as técnicas aplicadas frequentemente no universo digital em cada período de tempo, pautando-se nos ambientes metodológicos e padronizadamente construídos, como as bibliotecas digitais, por meio da catalogação e da indexação de recursos informacionais tradicionais, processos basilares do fazer bibliotecário.

Algumas dúvidas podem permear os usuários da *Web* quanto à sua origem e principalmente entender qual o motivo de tantas nomenclaturas adquiridas com o passar dos anos. Inicialmente trataremos do seu surgimento, que deve ser dissociado do nascimento da internet, considerada como uma,

[...] rede descentralizada de computadores interligados [...] que surgiu a partir da Arpanet, rede de segurança elaborada para fins militares nos EUA na década de 60, com o escopo de garantir a completa autonomia e independência entre os sistemas de envio de informações. (SAMPAIO; SOUZA, 2007),

A *WWW* foi criada posteriormente em 1989 por Tim Berners-Lee, pesquisador do CERN¹ neste período, ela foi originalmente concebida e desenvolvida para atender uma demanda de agilidade comunicacional, visando a partilha de informação automática entre os cientistas que trabalhavam em diferentes universidades e institutos de todo o mundo, sua ideia básica era fundir as

¹ CERN – é a Organização Europeia para a Pesquisa Nuclear, o nome é derivado do acrônimo para Conseil Européen pour la Recherche Nucléaire, ou Conselho Europeu de Pesquisa Nuclear.

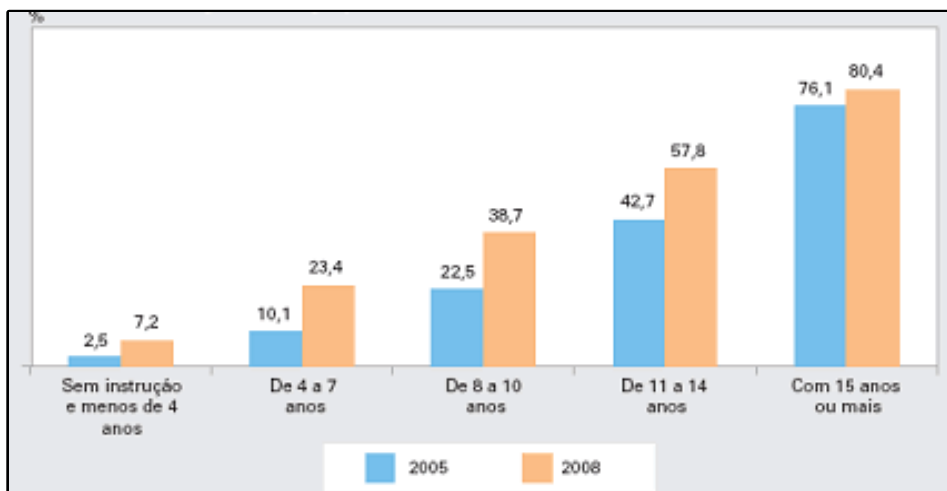
tecnologias de computadores pessoais, redes de computadores e hipertextos² em um poderoso sistema de informação e de fácil utilização global, ela é a “parte multimídia” da internet, possibilitando a exibição de páginas ou documentos que podem conter fotos, vídeos, textos, animações, sons e, sobretudo, permitir conexões com outros documentos através dos *links*.

Ao longo dos anos a *Web* foi ganhando adeptos e agregando recursos que potencializaram sua estrutura tornando-a responsável pela massificação do uso da internet, indo muito além dos objetivos pelos quais fora criada e por causa da grandiosa quantidade de usuários conquistada atraiu empreendimentos para o mundo virtual, tais como bancos, lojas e também bibliotecas, sendo que muitas empresas não existiam fisicamente, advindo exclusivamente do *ciberespaço*.

Para citar a força que esta rede de informação automática vem adquirindo exponencialmente no Brasil, pode-se observar uma pesquisa realizada com dados do Instituto Brasileiro de Geografia e Estatística- IBGE, a Pesquisa Nacional de Amostras por Domicílios - PNAD, realizada em 2008, referente ao *Acesso à Internet e Posse de Telefone Móvel Celular para Uso Pessoal*, esta traz diversos indicadores nos quais são apresentados um aumento significativo no número de usuários que acessaram a internet pelo menos uma vez nos últimos três meses anteriores. A pesquisa, até mesmo nas regiões Norte e Nordeste, que representavam o menor percentual brasileiro, na época, em relação a uma pesquisa semelhante realizada em 2005. Um destes indicadores se refere ao grau de escolaridade do internauta, à medida que se tem mais anos de estudo maior é o acesso “Entre aqueles com 15 anos ou mais de estudo, o percentual de usuários da rede era de 80,4%; entre os com 11 a 14 anos de estudo, 57,8%; com 8 a 10 anos de estudo, 38,7%; com 4 a 7 anos de estudo, 23,4%; e entre as pessoas sem instrução ou com menos de 4 anos, 7,2%” (IBGE, 2009). Todavia, nota-se nesta pesquisa que durante estes três anos houve um crescimento do acesso da população com menor escolaridade (menos de 15 anos de estudo), como se observa no gráfico 1:

² O hipertexto é um dos paradigmas básicos em que a teia mundial se baseia. Ele é uma espécie de texto multi-dimensional em que numa página trechos de texto se intercalam com referências a outras páginas.

Gráfico 1 - Comparativo percentual das pessoas que utilizaram a Internet nos últimos 3 meses, brasileiros, faixa etária acima dos 9 anos, segundo escolaridade – 2005/2008



Fonte: IBGE – PNAD 2005/2008

A estatística foi elaborada com base em pessoas na faixa etária a partir dos 10 anos, de ambos os sexos, que foram questionadas sobre o acesso à rede nos últimos três meses antes da pesquisa. O crescimento do acesso pode ser em função do barateamento nos preços dos computadores pessoais e dos provedores de acesso, facilitando o uso em casa e também pela disseminação do principal local apontado pelos entrevistados, os centros públicos de acesso pago, comumente chamados de *Lan Houses*”.

O motivo pelo qual as pessoas acessavam a internet era para “se comunicar com outras pessoas”, diferente do que aconteceu em 2005, onde o motivo principal era “educação ou aprendizado”, com isto foi possível observar o forte apelo da chamada “*Web Social*” ou “*Web 2.0*” e sua consolidação no decorrer destes três anos, pois diferente da chamada “*Web 1.0*”.

A *Web 2.0* possui como característica a interação do usuário com a informação e o surgimento das redes sociais, mecanismos extremamente importantes, mas que para um usuário mal orientado sobre práticas que possam favorecer seu crescimento intelectual, retratará um subaproveitamento de instrumentos poderosos de comunicação contidos na *Web* que poderiam ser melhor captados para o desenvolvimento educacional.

Neste ponto é necessária a presença de um bibliotecário neste ambiente. Estar presente em *sites*, nas redes sociais, levando a biblioteca àqueles que talvez nunca tenham estado em uma, ou seja, esta pesquisa reforça a existência de uma

demanda de usuários na rede, ávidos por comunicação e interação abrindo precedentes para a atuação dos profissionais da informação que possam implantar, por exemplo bibliotecas digitais, aproveitando-se da interatividade da *Web 2.0* aliando a performance prometida pela *Web Semântica*³, para transformar a ideia de biblioteca como um lugar que une a interação social, aprendizado e modernidade.

Foi perceptível na análise da pesquisa realizada pelo IBGE, citações de algumas nomenclaturas diferentes atribuídas à *Web*, estas terminologias acabaram por representar o desenvolvimento desta rede de documentos que no transcorrer da sua história conseguiu adquirir características mais versáteis para transferência, exibição e interação de informações multimídia.

2.1 A evolução da *Web*

Segundo Blattman e Silva (2007, p. 2) “A evolução da *web* possibilita a criação de espaços cada vez mais interativos, nos quais os usuários possam modificar conteúdos e criar novos ambientes hipertextuais”. De acordo com Oliveira (2008, p. 14), “Essa evolução pode ser considerada como “eras”, dando origem ao que muitos chamam de ‘*web eras*’”.

Nesse contexto nota-se que à medida que a *Web* evoluiu ganhou uma espécie de novo nome ou nova versão, o que acontece frequentemente com *softwares* quando agregam novas funcionalidades, valendo ressaltar que não são construídos novos programas e sim adicionadas tecnologias que possibilitam um melhor desempenho do mesmo.

Um conceito semelhante pode ser aplicado a esta evolução da *WWW*, conforme Guimarães (2009, p. 3), citando uma entrevista do criador da *Web*, Tim Berners-Lee para um *podcast*⁴ do *site* da IBM sobre a versão 2.0 da *Web*, se esta era muito diferente da primeira geração, obtendo a seguinte resposta:

³ A *Web Semântica* surge como um conjunto de práticas estruturadas em camadas de desenvolvimento propondo a organização no ambiente virtual através de padrões de metadados, proporcionando intercâmbio e interoperabilidade de dados além de associar significado aos mesmos por meio de ontologias visando uma recuperação mais precisa das informações. (Este termo será mais explorado na seção 4 desta pesquisa).

⁴ Podcast é uma forma de transmissão de arquivos multimídia na Internet criados pelos próprios usuários. Nestes arquivos, as pessoas disponibilizam listas e seleções de músicas ou simplesmente falam e expõem suas opiniões sobre os mais diversos assuntos, como política ou o capítulo da novela. Pense no podcast como um blog, só que ao invés de escrever, as pessoas falam. Disponível

Totalmente não. A *Web 1.0* foi toda para conectar as pessoas. Era um espaço interativo, e eu acho a *Web 2.0*, naturalmente, um jargão, ninguém sabe mesmo o que significa. Se a *Web 2.0* para você são blogs e wikis, então, isto varia de pessoa para pessoa. Mas isso foi o que a *Web* era para ser, tudo junto. E, na verdade, você sabe, esta *Web 2.0*, significa usar as normas que tenham sido produzidas por todas essas pessoas trabalhando em *Web 1.0*. (GUIMARÃES, 2009, p. 3 *apud* LANINGHAN, 2006).⁵

Conforme o que foi dito por Tim Berners-Lee, a transição dos termos *Web 1.0* para *Web 2.0*, foram apenas jargões. No momento em que emergiam comentários sobre esta segunda versão ninguém sabia ao certo o que ela significava, pois em tese, a primeira versão também foi criada para conectar pessoas, embora nesta fase tal procedimento tenha se realizado de forma mais rudimentar e tudo ainda era novidade; quem mais se aproveitou deste período foram as empresas que podiam expor seus produtos para uma quantidade muito maior de clientes, entretanto, havia pouquíssima interação dos clientes com as empresas ou dos clientes entre si. O que foi modificado posteriormente através da inserção de comentários nos *sites*, nascimento dos blogs, redes sociais, entre outros elementos atribuídos a uma “nova” geração da *Web*, que transferiu o foco de desenvolvimento, da empresa para o usuário, ou seja, para a interação social. A partir deste momento, mais pesquisadores passaram a identificar e agrupar essas características mais recentes; assim como em alguns *softwares*, criou-se a denominação *Web 2.0*.

No entanto, a diferença essencial entre a *Web 1.0* e a *Web 2.0* é que os criadores de conteúdos eram poucos na *Web 1.0*, com a grande maioria dos usuários simplesmente agindo como consumidores de conteúdo, enquanto que qualquer participante pode ser um criador de conteúdo na *Web 2.0* e numerosos auxílios tecnológicos têm sido elaborados para maximizar o potencial de criação de conteúdo. (CORMODE; KRISHNAMURTHAY, 2008)⁶

em: <<http://www.tecmundo.com.br/1252-o-que-e-podcast-.htm#ixzz2QZcKNEUE>>. Acesso em 20 fev. 2013.

⁵ Tradução Livre de: *Totally not. Web 1.0 was all about connecting people. It was an interactive space, and i think web 2.0 is of course a piece of jargon, nobody even knows what it means. If Web 2.0 for you is blogs and wikis, then that is people to people. But that was what the Web was supposed to be all along. And in fact, you know, this Web 2.0, it means using the standards which have been produced by all these people working on web 1.0.*

⁶ Tradução Livre de: *However, the essential difference between Web 1.0 and Web 2.0 is that content creators were few in Web 1.0 with the vast majority of users simply acting as consumers of content, while any participant can be a content creator in Web 2.0 and numerous technological aids have been created to maximize the potential for content creation.*

Entretanto, a discussão sobre este período de transição, que fez surgir oficialmente uma nova terminologia, ainda persiste, isto aconteceu em outubro de 2004 em sessões de *Brainstorming* ocorridas nas conferências realizadas pela *O'Reilly Media, Inc.* e a *MediaLive International*, empresas que promovem eventos e produzem conteúdos voltados para a área de tecnologia da informação; muitos autores demonstram uma certa prudência em relação a esta mudança.

Em um ano e meio desde que surgiu, o termo “*Web 2.0*” tem claramente tomado conta, com mais de 9,5 milhões de citações no Google (definição publicada na *Web* em Setembro de 2005; 135 milhões de citações em Fevereiro de 2007). Mas há ainda uma enorme quantidade de desacordo sobre o que significa *Web 2.0*, com algumas pessoas condenando-o como um chavão de marketing sem sentido, e outros a aceitá-la como a nova sabedoria convencional. (O'REILLY, 2007, p.18)

O que pode ser confirmado nesta colocação de Corrêa (2008, p. 7) que indaga “A suspeita principal recai em saber se *Web 2.0* delineia uma nova concepção associada a inovações tecnológicas para aumentar a participação do usuário final ou se representa apenas mais uma palavra da moda, funcionando como marketing *buzzword*?”,

Primo (2007, p. 1), diz que

A *Web 2.0* refere-se não apenas a uma combinação de técnicas informáticas (serviços *Web*, linguagem Ajax, *Web syndication* etc.), mas também a um determinado período tecnológico, a um conjunto de novas estratégias mercadológicas e a processos de comunicação mediados pelo computador.

O que se pode entender é que independente de nomeações para esta sub-rede da internet, novas formas de desenvolvimento foram elaboradas para comportar formatos mais interativos de documentos.

Com o apoio da base bem sucedida da versão anterior quanto à disseminação da ascendente produção de recursos informacionais, em pouco tempo notou-se que o conteúdo não poderia permanecer estático, em nome de um melhor desempenho de produções acadêmicas e atração de novos clientes, alguns setores resolveram investir em práticas que pudessem favorecer esta ideia e foram justamente as empresas que reuniam características específicas que conseguiram sobreviver ao chamado estouro da bolha pontocom.

O conhecimento já não é uma pirâmide estática, cresce e se movimenta por uma extensa rede móvel de laboratórios, de centros de pesquisa, de bibliotecas, de bancos de dados, de procedimentos técnicos, de meios de comunicação, dispositivos de gravação e rede

de medição contínua a se espalhar no mesmo movimento entre humanos e não-humanos, associando moléculas e grupos sociais, elétrons e instituições.⁷.(LÉVY, 2004, p. 127).

Retornando ao artigo publicado por Berners-Lee, em 2007, quando já havia um certo senso comum e milhões de citações do nome “*Web 2.0*”, ele coloca o termo novamente em pauta para tentar explicar que houve algumas características semelhantes em determinados *sites* de empresas que sobreviveram ao estouro da bolha pontocom, que apontavam um surgimento de uma forma mais eficiente de se desenvolver aplicativos, contudo, estes supostos requisitos não são absolutamente rígidos principalmente se o enxergarmos como uma plataforma.

Como muitos conceitos importantes, a *Web 2.0* não tem um limite rígido, mas em vez disso, um núcleo gravitacional. Você pode visualizar a *Web 2.0* como um conjunto de princípios e práticas que unem um verdadeiro sistema solar de sites que demonstram alguns ou todos esses princípios, a uma distância que varia do núcleo. (2007, p. 18-19)⁸

Em meio a tanta especulação sobre a primeira e a segunda geração da *Web* tentaremos demonstrar aqui, quais as características atribuídas a cada versão que mais representam os respectivos momentos da rede listando e trazendo exemplos práticos de aplicações *Web* voltadas a área da biblioteconomia.

2.1 Características atribuídas as *Web 1.0* e *Web 2.0*

A *Web 1.0* foi a primeira versão da *World Wide Web* e trouxe ao ambiente virtual, facilidades na transmissão de documentos, popularizando uma maneira diferente de publicar e obter informações. Fazendo uso de hipertextos, *links* e como já foi mencionado, suas páginas eram estáticas e conteúdo estratificado baseado na estrutura taxonômica, ou seja, estrutura hierárquica, onde o usuário não podia editar as publicações.

⁷ Tradução livre de: *El conocimiento ya no es una pirámide estática, crece y se desplaza por una extensa red móvil de laboratorios, de centros de investigaciones, de bibliotecas, de bancos de datos, de procedimientos técnicos, de medios masivos, de dispositivos de registro y de medida, red que no cesa de extenderse con el mismo movimiento entre los humanos y los no humanos, asociando moléculas y grupos sociales, electrones e instituciones.*

⁸ Tradução do artigo de Berners-Lee

De acordo com Coutinho e Betenttuit Junior (2008, p.1861), a *Web* neste período,

[...] era bastante onerosa para os seus utilizadores; a grande maioria dos serviços eram pagos e controlados através de licenças, os sistemas eram restritos a quem detinha poder de compra para custear as transacções online e adquirir o software para criação e manutenção de sites.

Quanto as tecnologias podemos observar o que foi utilizado como base de desenvolvimento desta rede, Tim Berners-Lee em 1989, criou e uniu três simples tecnologias, segundo consta no *site* do W3C Brasil⁹:

- Localizador ou Identificador uniforme de Recursos (URI ou URL) - para identificar *recursos* (por exemplo, documentos, dados) *na Web*, e saber onde encontrá-los;
- Hypertext Markup Language (HTML) - linguagem para representar o conteúdo em termos de páginas na *Web* e de expressá-las em links;
- Hypertext Transfer Protocol (HTTP) - protocolo para mover dados na *Web* e em toda a Internet;

Além disso, um dos grandes motivos para a popularização da *Web* dependeria de um quarto fator importante, destacada pela tecnologia aberta e disponibilizada a todos de forma gratuita.

Da *Web* 1.0 para a *Web* 2.0 o direcionamento mudou, pois se na primeira geração da *Web* procurava-se adaptar as pessoas a este novo paradigma de comunicação através de navegadores gráficos, a quebra da sequência lógica única para ler um documento com o advento dos *hiperlinks* entre outras peculiaridades do mundo digital, na segunda geração fica clara a consolidação da *Web* e o constante crescimento do número de adeptos com 2 bilhões de pessoas que acessam, 100 milhões de sites no ar. (de acordo com o W3C – *World Wide Web Consortium*¹⁰). Como simplifica Davis (2005) a *Web* 1.0 direcionava pessoas para a informação e a *Web* 2.0 intenciona levar a informação para as pessoas.

Nesta versão que vigora até os dias atuais direcionou-se a atenção para os usuários, as aplicações ficaram mais interativas, *sites* permitem que as pessoas não apenas vejam as suas notícias mas também opinem sobre elas; o conteúdo é

⁹ <http://www.w3c.br>. Acessado em 28 dez. 2012.

¹⁰ *idem*. Acessado em 12 jan. 2013.

dinâmico, são atualizados constantemente, *sites* pessoais foram substituídos por blogs, dispensando em sua grande maioria a presença de um profissional de informática para sua construção bastando apenas um conhecimento básico em informática para o manejo das tecnologias da plataforma a qual se pretende desenvolver, pois a estrutura já está previamente formatada, abrindo espaço para modificação em HTML, uso de Cascading Style Sheets - CSS , adição ou remoção de mídias, caso esteja se tratando de um usuário mais avançado que prefira trabalhar com codificação ao invés da interface gráfica.

Uma variação dos *Weblogs* que tem se popularizado ultimamente são os chamados *Vlogs*, com o mesmo propósito dos primeiros, que é expor opiniões pessoais sobre questões diversas ou sobre a própria vida do seu criador, mas por meio de gravações de vídeos que serão socializadas e comentadas na *Web*, ajudam a definir o caráter social da *Web* Sintática (ou *Web* Social).

O grande marco da *Web* Social é a inteligência colaborativa, onde o conhecimento é construído pelos usuários para os usuários, os mesmos “postam”, colaboram com informações a respeito de determinado assunto, enquanto outros complementam ou mesmo corrige-os quando for o caso, produzindo um conhecimento uno, advindo da parceria de muitas pessoas. Um bom exemplo disto é a *Wikipedia*; uma enciclopédia livre, em rede e colaborativa, onde o conhecimento é construído pelos próprios usuários, outro exemplo característico desta fase é o desenvolvimento concentrado em esforços folksonômicos, ou seja, as pessoas empregam etiquetas em linguagem natural, as chamadas *tags*, nos conteúdos que pretendem evidenciar, por meios de palavras-chave que os identifique. Esta indexação acontece de modo *bottom-up* , (em outros termos: do específico para o geral) e é amplamente aplicada em contextos sociais participativos formando uma nuvem de indicadores para agregar informações similares com o propósito de permitir a localização de recursos informacionais digitais.

Ainda baseada nas declarações de Berners-Lee, Hendler e Lassila (2001) podemos extrair dez características predominantes na chamada *Web* Social, Sintática ou 2.0:

A *Web* é uma plataforma. Nós passamos de software instalado em seu PC para serviços baseados em software acessíveis online. Todos os dados e software estão agora disponíveis desta forma.

A Web é funcionalidade. A *Web* ajuda na transferência de informações e serviços dos sites.

A Web é simples. Facilita o acesso e o uso dos serviços da *internet* usando interfaces amigáveis.

A Web é clareza. Os modelos de desenvolvimento, os processos e os modelos de negócio tornam-se transparentes. A clareza é associada com a capacidade de compartilhar informações e serviços com facilidade, e tornar possível através da implementação de elementos modulares intuitivos.

A Web é social. As pessoas criam a *Web*, “juntam-se a *Web*”, através da socialização e da transferência de membros do mundo físico para o mundo *online*.

A Web é atualização. Os usuários são vistos como co-desenvolvedores, enquanto a *Web 2.0* continua em “beta permanente”, onde permanece em fase de desenvolvimento por um período indefinido de tempo.

A Web é diversidade. Esta ideia é semelhante ao conceito da Cauda longa, que foca em conteúdos menos populares, que não podiam ser acessados em ambientes convencionais.

A Web é miscível. A expansão dos códigos, a fim de modificar os aplicativos *Web* permite que indivíduos, não necessariamente profissionais de informática, possam misturar diferentes aplicações, a fim de criar novas. A *Web 2.0* obtém seu poder através deste “*mashup*”¹¹ de capacidade.

A Web é participativa. A *Web 2.0* adotou uma estrutura de participação que incentiva os usuários a melhorar o aplicativo enquanto eles usam, em vez de mantê-lo rígido e controlado.

A Web está em nossas mãos. Com a sua organização crescente, a caracterização de informações enfatiza a fácil interação com o usuário através de um aprofundamento de *links*, graças a fenômenos como “rotulagem/etiquetagem” social, as informações estão cada vez mais facilmente disponíveis.

Para melhor ilustrar as diferenças tecnológicas das práticas usuais em cada variante da rede, mostraremos um comparativo no quadro 1, onde são listadas práticas e sites com características da *Web 1.0* (lado esquerdo) e do lado direito os da *Web 2.0*:

¹¹ Segundo a Wikipedia: Um mashup é um site personalizado ou uma aplicação web que usa conteúdo de mais de uma fonte para criar um novo serviço completo.

Quadro 1 – Comparativo entre tecnologias associadas à *Web 1.0* e *2.0*

WEB 1.0	WEB 2.0
Ofoto Mp3.com Britannica Online Sites Pessoais Publicar Sistemas Fechados Taxonomia (diretório)	Flickr Napster Wikipedia Blogs Participar Wikis Folksonomia (tagging)

Fonte: Blattmann e Silva (2007, p.198)

Vemos com isto que algumas abordagens e aplicativos da *Web 1.0* demonstrados no quadro acima tem seus “equivalentes” mais colaborativos, como a Wikipedia em relação a Britannica Online, facilidades dos blogs em comparação aos sites pessoais, mudança das “palavras-chave” de publicar (1.0) para participar (2.0) e dentre outros a estrutura em diretórios hierárquicos (taxonomia) para a indexação em linguagem natural com tags (marcadores), definidos pelos usuários (folksonomia), ratificando assim a presença do usuário como o foco da segunda geração da *Web* que neste quadro fica claro através destes exemplos no comparativo com a geração anterior .

2.2 A Biblioteconomia no contexto *Web*

A profissão de bibliotecário ao contrário de algumas previsões de extinção se torna cada vez mais necessária para a sociedade do século XXI, à medida que crescem vertiginosamente a produção e o consumo de informações, a presença de um especialista capacitado para organizar, disseminar e recuperar recursos informacionais conforme os anseios dos usuários, viabiliza uma solução para pôr ordem, em especial, ao caótico ambiente *Web* que carece de um profissional com tal perfil.

Para tanto, além das já conhecidas bibliotecas tradicionais, centros de documentação e outros campos de atuação clássicos, os bibliotecários da

atualidade precisam assumir o seu espaço no ambiente virtual, trabalhar a interdisciplinaridade em favor do fortalecimento e expansão das suas áreas de atuação, traçando, quando necessário, uma parceria com a comunidade da Ciência da Computação visando desenvolver as tecnologias emergentes e consolidar as que já são uma realidade aplicável.

A Ciência da Informação e a Ciência da Computação têm bastante em comum, a exemplo de serem capazes de fornecer pilares para a arquitetura da informação, ficando claro que os estudos de cada comunidade dará ênfase ao seu campo específico, entretanto, agindo em parceria complementam-se corroborando, por exemplo, para a descrição e a representação dos dados na *Web*, o que resulta na melhoria da recuperação de documentos para os usuários finais, que se beneficiarão desta colaboração mútua recebendo resultados mais precisos em suas buscas *online*.

A Biblioteconomia também esteve presente no desenvolver da *Web*, através de sites da área, blogs catálogos online, dentre outros, assim como outras áreas, embora que de modo mais tímido que as empresas que tinham interesses comerciais fazendo uso dos *Web Services*¹² com maior desenvoltura, ou seja, explorando um maior quantitativo de práticas e tecnologias referentes ao espaço cibernético.

Visando atrair mais pessoas, este ramo da Biblioteconomia e da Ciência da Informação começou a aparecer através de sites 1.0 para divulgar e promover bibliotecas “físicas”, passando pela construção das ainda um tanto enigmáticas bibliotecas 2.0, chegando por fim às bibliotecas semânticas, estas últimas serão mais exploradas nos capítulos subsequentes.

Abram (2007) sugere uma lista de 23 exercícios realizados em mais ou menos nove semanas para que os bibliotecários estejam inteirados e possam expandir seu conhecimento sobre a *Web 2.0*, dentre eles construir blogs, conhecer feeds RSS e configurá-los, conhecer o *Del.icio.us*¹³ com o intuito de assimilar mais sobre *tags* e folksonomia, conhecer o instrumento de busca por blogs *Technorati* e aprender como trabalhar com marcas de posts, ler sobre algumas perspectivas referentes a *Web 2.0*, Biblioteca 2.0 e o futuro das bibliotecas, saber mais sobre os

¹² Entenda-se *Web Services*, neste caso, como serviços *Web* em geral.

¹³ *Del.icio.us*: *site* que oferece serviços *online* permitindo adicionar e pesquisar *bookmarks* sobre qualquer assunto.

Wikis além de descobrir algumas maneiras inovadoras para que as bibliotecas possam usá-los.

Nesta pesquisa não iremos explorar detalhadamente cada tecnologia empregada na *Web 2.0*, apenas mencionarmos o que já foi, vem e pode ser empregado na área da Biblioteconomia, para que possa ser incorporado às bibliotecas digitais.

Blattman e Silva (2007) sugerem um comparativo entre os instrumentos usados nas chamadas biblioteca 1.0 e biblioteca 2.0 que pode ser observado no quadro 2:

Quadro 2 - Evolução da Biblioteca 1.0 para Biblioteca 2.0

Biblioteca 1.0	Biblioteca 2.0
<input type="checkbox"/> Correio eletrônico e páginas de questões mais frequentes (FAQ)	<input type="checkbox"/> Serviço de referência via bate-papo (Chat)
<input type="checkbox"/> Tutorial baseado em texto	<input type="checkbox"/> Mídia interativa (Streaming media) em bases de dados
<input type="checkbox"/> Listas de correio eletrônico, Webmasters	<input type="checkbox"/> Blogs, Wikis, Leitoras de RSS
<input type="checkbox"/> Esquemas de classificação controlada	<input type="checkbox"/> Indexação com base em esquemas controlados
<input type="checkbox"/> Catálogo impresso	<input type="checkbox"/> Catálogo com agregados, blogs, wikis e páginas Web

Fonte: Blattman e Silva (2007, p.196)

Em relação às bibliotecas digitais, verificamos que seu conceito está em constante reformulação, segundo Sayão (2009, p. 16),

O que se pode concluir com algum risco é que o conceito de biblioteca digital é algo ainda no estágio transiente de evolução e que provavelmente guardará significados distintos ou receberá denominações distintas à medida que as atuais sobreposições conceituais se resolvam.

Mas assumimos aqui, este ponto de vista de Alvarenga (2001, p. 3)

De forma simplificada, a biblioteca digital, pode ser definida como um conjunto de objetos digitais construídos a partir do uso de instrumentos eletrônicos, concebidos com o objetivo de registrar e comunicar pensamentos, idéias, imagens e sons, disponíveis a um contingente ilimitado de pessoas, dispersas onde quer que a plataforma *www* alcance.

Diante de tal conceito, dentre tantos espaços para atuação do profissional da informação na *WWW* encontramos neste, um bom exemplo de onde os bibliotecários podem desenvolver as habilidades que já possuem para a organização de documentos impressos, pois o digital é apenas um suporte que propicia maior agilidade para disseminar a informação, nada que represente uma antítese ao trabalho exercido pelo bibliotecário, ambas são tecnologias e cada uma predomina num dado momento da história da comunicação e da evolução dos suportes dos registros do conhecimento.

Deste modo são complementares e não adversárias devendo o profissional da informação fazer uso das tecnologias que melhor atenda as suas condições de trabalho, mas ter uma noção geral das mais diversas formas tecnológicas sem opor resistência, essencialmente às que vem emergindo com a popularização da informática.

O digital não é o antagonico do impresso, como o rolo de papiro não é o antagonico do livro. Para cumprir o seu papel ancestral a biblioteca sempre se apropriou das mais avançadas tecnologias disponíveis e vem continuamente evoluindo no ritmo dessas tecnologias. Assim foi com a tecnologia de microfilme, com a computação e agora com a *Web*. (SAYÃO, 2009, p.16).

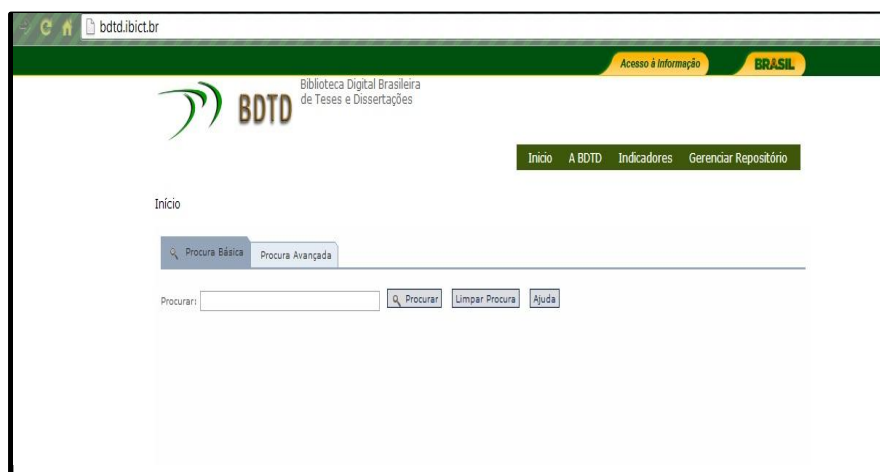
Existem relativamente poucas bibliotecas que se utilizam das práticas 2.0 e quase nenhuma biblioteca semântica, isto serve para ratificar a importância de estudos por profissionais da Biblioteconomia em conhecer e consolidar estas técnicas proporcionando uma maior visibilidade deste tipo de bibliotecas que não são muito conhecidas, até mesmo no meio acadêmico. Vejamos alguns exemplos de bibliotecas digitais:

Figura 1 – Página Inicial da Biblioteca Digital Mundial



Fonte : <http://www.wdl.org/pt/> . Acesso em: 18 de jan. 2013

Figura 2 – Página Inicial da Biblioteca Digital Brasileira de Teses e Dissertações – BDTD



Fonte: <http://bdtd.ibict.br/>. Acesso em: 18 de jan. 2013

Ao compararmos as duas figuras, podemos presenciar o maior apelo da Web Social na Figura 1 referente à Biblioteca Digital Mundial, logo na sua página inicial visualizamos uma lista com as atualizações da biblioteca, na parte inferior tem um botão que dá acesso ao microblog *Twitter*, para que os usuários possam segui-la interagindo entre si e com a biblioteca, esta que abre um canal de comunicação com

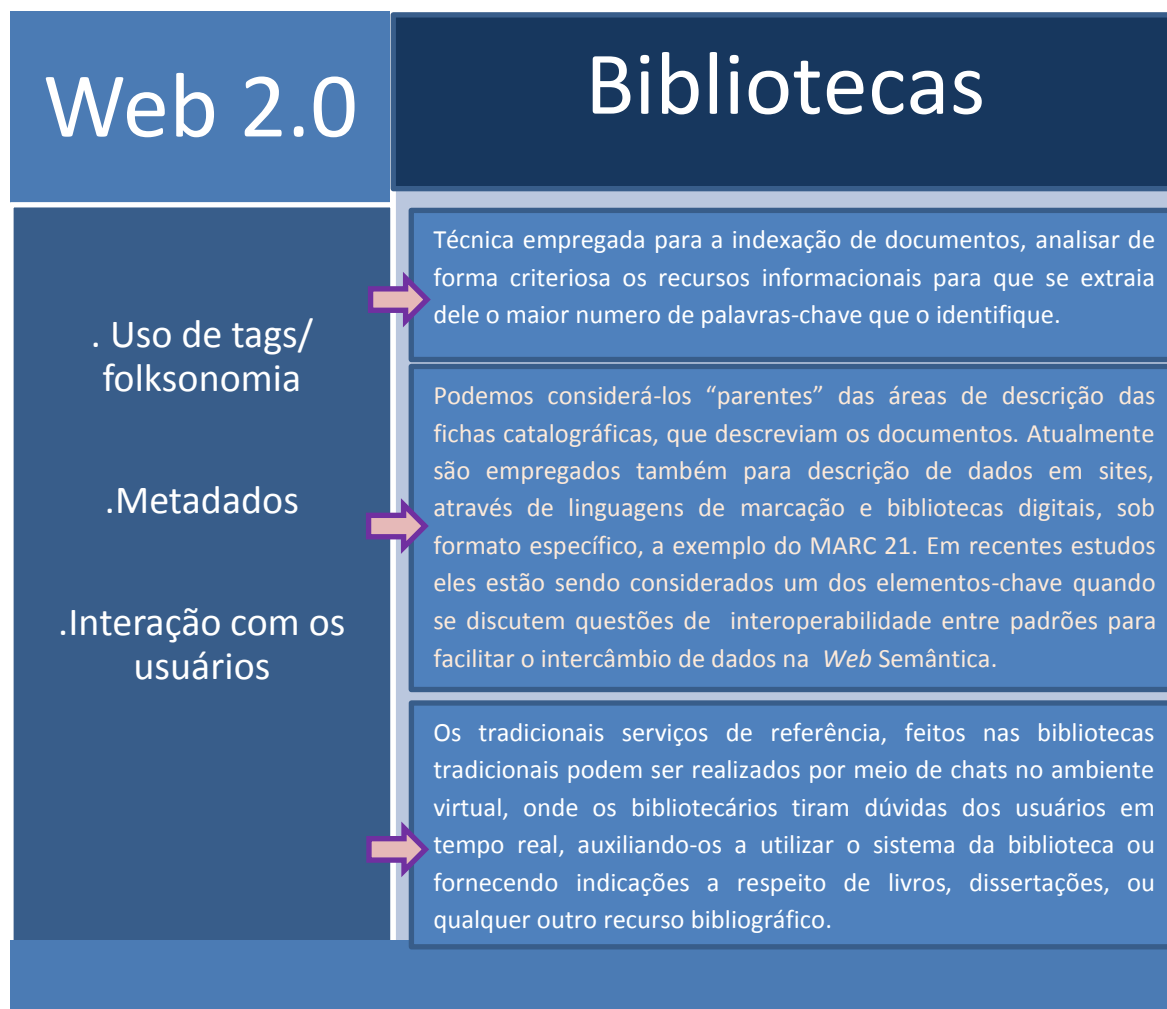
os usuários sendo que os posts atualizados do *Twitter* também são mostrados na *homepage*, na lateral esquerda da página, além do uso de gráficos e figuras.

Já na página inicial da Biblioteca Digital Brasileira de Teses e Dissertações não é tão aparente tais apelos, pois como pode ser observado na figura 2, ela possui uma “interface limpa” lembrando a página do Google, o que é eficaz para a consulta, mas pode não ser tão eficiente quanto à atração e à interação de usuários.¹⁴

Nesta pequena comparação vimos que não há uma etapa cronológica tão definida quanto ao uso de tecnologias 1.0 e 2.0, assim como pode existir aplicativos *Web* que se utilizem de “tecnologia de ponta” existem os que fazem uso de recursos 2.0 em menor escala ou os que tenham características da primeira geração; o que vale neste momento é o bom senso e a análise do público-alvo para a qual a aplicação se destina; contudo, deve-se desenvolver este pensamento crítico com a finalidade de extrair o melhor de cada período da *Web* (ou instrumento da *Web*).

Tomando como base a *Web* predominante (2.0) veremos como algumas características implantadas no ambiente virtual já são utilizadas pelos bibliotecários, podendo fazer um paralelo, conforme apresentado no quadro 3:

¹⁴ Nesta pesquisa apontamos unicamente a página inicial das bibliotecas, sem maior aprofundamento sobre possíveis trabalhos de divulgação em redes sociais.

Figura 3 – Comparativo entre serviços da *Web 2.0* e serviços da biblioteca tradicional

Fonte: Elaborado pela autora.

Como foi exemplificado anteriormente, a segunda geração da *WWW* veio trazer uma revolução para as comunicações humanas que se tornaram mais ágeis e eficientes, a atenção volta-se para o usuário e isto se reflete também no seu desenvolvimento, o uso de linguagens naturais facilitou a interação entre os povos, favoreceu a expansão de publicações e a criação de conteúdos, tornando a sociedade gradativamente dependente de tecnologia; todavia, as formas de comunicação humanas não são compreensíveis para as máquinas que estão no intermédio deste processo gerando muitos equívocos aos agentes que efetuam buscas de informações automáticas.

Neste cenário surge a proposta da *Web Semântica*, visando sanar os problemas na recuperação de informação por meio da catalogação na *Web* através

de descrição do conteúdo de forma semântica auxiliando a identificação dos documentos neste ambiente, que será discutido no capítulo 4.

2.3 A catalogação e suas contribuições para a Web

A razão da existência das bibliotecas e seus funcionários é atender aos seus usuários para que os mesmos tenham acesso à informação desejada, para tanto, o bibliotecário precisa preparar os recursos bibliográficos de tal forma que sejam acessíveis ao seu público, e isto é facilitado com a representação descritiva dos documentos.

Com o intuito de que os usuários consigam encontrar sobre a informação requerida com maior precisão, independentemente do seu suporte, uma descrição bem feita (quanto mais detalhada melhor) de um documento, trará agilidade à sua recuperação e localização, cumprindo primordialmente uma das cinco Leis da Biblioteconomia¹⁵ propostas por Ranganathan, que é poupar o tempo do leitor.

A catalogação é um processo de representação descritiva dos documentos que procura detalhá-los ao máximo quanto à sua forma e ao seu conteúdo, através de códigos e padrões da qual resultam os catálogos, servindo para consulta dos usuários, promovendo a recuperação e identificação do recurso bibliográfico.

Existem vários tipos de catálogos e de acordo com o suporte no qual estão inseridos mudam as terminologias e códigos empregados, por exemplo, na *Web* está bastante comum o emprego do termo “metadados”.¹⁶

Entretanto, alguns pesquisadores fora do campo da Biblioteconomia ainda encaram este tema como novidade, diferente de Breitman (2005, p.12) que reconhece que,

[...] o conceito já é aplicado por bibliotecários, museólogos e arquivistas há muito tempo e ressalta a importância da disciplina de Catalogação dentro Ciência da Informação, algo que não ocorria,

¹⁵As referidas Leis são:

1. Os livros são para usar.
2. A cada leitor seu livro.
3. A cada livro seu leitor.
4. Poupe o tempo do leitor.
5. A biblioteca é um organismo em crescimento.

¹⁶ Metadados são um conjunto de dados-atributos, devidamente estruturados e codificados, com base em padrões internacionais, para representar informações de um recurso informacional em meio digital ou não – digital, contendo uma série de características e objetivos. (ROSETTO, 2003, p.10).

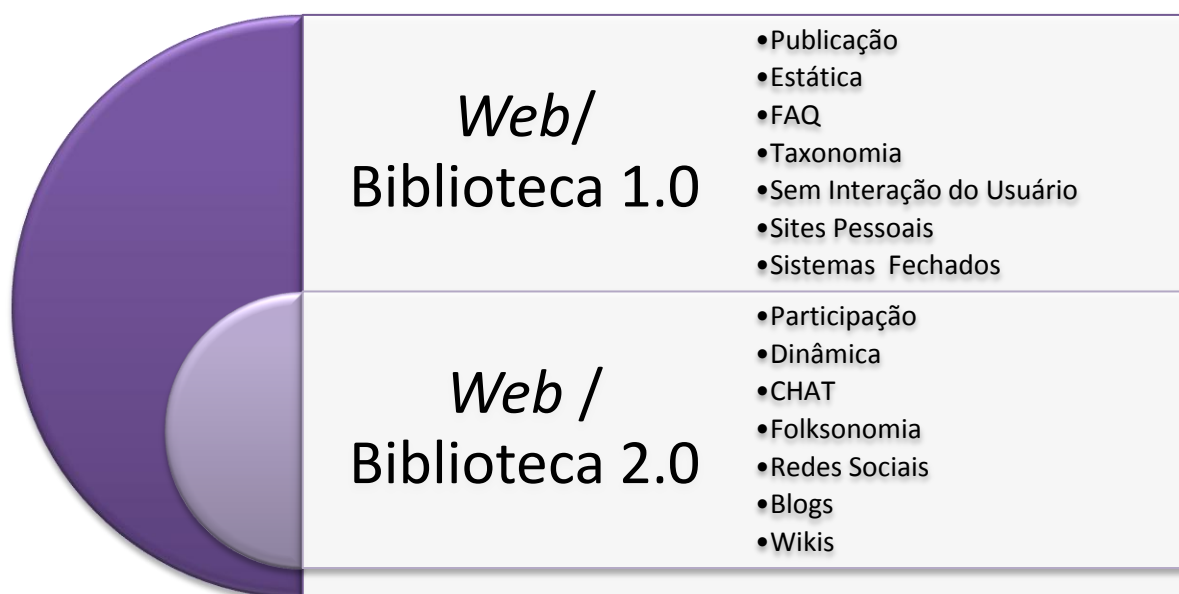
pois a matéria estava sendo considerada um tanto arcana e presa ao manuscrito.

Como podemos verificar isto não condiz com a realidade atual, o estudo dos metadados são essenciais principalmente para a catalogação na *Web* como um todo. No que tange à chamada *Web 3.0* a análise se estende por padrões e *frameworks* que possam organizar e auxiliar numa futura unificação de padrões de metadados visando à interoperabilidade, para agilizar o retorno de buscas com solicitações cada vez mais complexas e que estas sejam compreensíveis por máquinas em escala mundial.

Seguindo a linha de raciocínio utilizada pela Biblioteconomia ao longo do tempo, que tem como base a descrição de recursos informacionais através de metadados padronizados objetivando a inferência do conteúdo e o fácil acesso aos documentos, podemos ratificar a importância de levar esses conceitos para as bibliotecas digitais e, a partir destas, traçar um paralelo com o que se tem proposto no projeto *Web Semântica*, nas suas três camadas-base, tais como, URI/URL, XML e RDF. Estas têm também por finalidade a descrição, a identificação, a padronização e a interoperabilidade dos dados, em conjunto com outras tecnologias usadas em camadas superiores pretendem organizar a *Web*, descrevendo o conteúdo dos documentos através de metadados, linguagens de marcação de texto e ontologias, atribuindo uma relação contextual para “*linkar*” as informações das bases de dados, estas que deverão estar padronizadas para que possam “conversar” entre si, visando recuperar informações mais precisas.

Com base no exposto entre *Web 1.0* e *2.0*, bem como seus conceitos, tecnologias e evolução, articulados às bibliotecas, a figura 3 destaca elementos comuns potencialmente aplicados nesses ambientes informacionais.

Figura 4: Características comuns simultaneamente à Web e Bibliotecas 1.0 e 2.0



Fonte: Elaborado pela autora.

O que propomos neste capítulo foi elaborar uma breve introdução ao mundo da *Web*, mostrar de forma sucinta como ela evoluiu, as discussões sobre suas nomenclaturas e deixar claro que cada geração tem suas características marcantes; todavia, uma não se sobrepõe totalmente a outra, ou seja, mesmo estando ainda no período da *Web* Sintática ou *Web* 2.0 não quer dizer que não tenhamos *sites* com características da *Web* antecessora ou até mesmo da sucessora, *Web* Semântica.

Foi mostrado também exemplos destas características em bibliotecas digitais nacionais e internacionais. Desse modo, trouxemos a Biblioteconomia para o ambiente virtual por este viés envolvendo também a catalogação e os metadados, conceito fundamental para o intercâmbio de dados no meio digital e que vem sendo estudado pelos vários profissionais da informação, por ser uma das “peças-chave” para implementação da *Web* Semântica.

Assim, demonstra-se a necessidade do profissional bibliotecário estar inserido neste ambiente, para que se desenvolva cada vez mais a capacidade de atender aos usuários em todos os ambientes e que se possa compreender que a *Web* Semântica não surgiu ao acaso, e sim de uma transformação gradual, aperfeiçoamento e observação de tecnologias utilizadas na área de Biblioteconomia.

A partir deste conjunto de instrumentos tecnológicos podemos extrair boas práticas para as nossas bibliotecas digitais, objeto de estudo do capítulo 3.

3 BIBLIOTECAS DIGITAIS: CONTEXTO, CONCEITOS E TECNOLOGIAS

Vimos no capítulo anterior, um resumo sobre a evolução da *Web* e como as bibliotecas também vem conquistando seu espaço gradativamente dentro deste contexto, se fazendo presentes através de *blogs*, *Websites*, redes sociais e também das bibliotecas digitais, trazendo para este ambiente a maioria dos serviços oferecidos na biblioteca tradicional, acrescidos das facilidades próprias do meio cibernético.

Neste capítulo trataremos um pouco mais sobre as bibliotecas digitais, estas, que são foco de diversas discussões acadêmicas gerando controvérsias sobre sua definição, mostrando que elas são resultados de um processo evolutivo que está em constante transformação e, ainda, há muito que contribuir para a organização informacional na *Web* e para a modernização da Biblioteconomia.

Parafraseando o célebre químico Antoine Lavoisier “Na natureza nada se cria, nada se perde, tudo se transforma”. Podemos assim iniciar este breve estudo sobre as bibliotecas digitais, entendendo que a preocupação com o registro, disseminação e armazenamento de informações, questões essenciais para a biblioteca digital, não é uma novidade, vem antes mesmo da invenção da imprensa com Gutenberg, em 1455, estas questões nos remetem à pré-história, onde o ser humano começa a registrar os seus conhecimentos.

As pinturas rupestres segundo pesquisadores da área, representavam o cotidiano dos habitantes das cavernas, não se sabe o porquê destes registros mas eles acabaram se tornando um meio de transmissão, do que era vivenciado naqueles tempos, mais confiável do que a memória e mais duradouro que a comunicação oral. “As hipóteses quanto aos sentidos e aos objetivos dessas imagens variam de acordo com a época em que foram estudadas e a perspectiva teórica adotada.” (NOVAES, 2008, p. 458).

Percebemos que a partir deste período remoto até a contemporaneidade, a ideia central é que possamos registrar nossos conhecimentos em um lugar seguro, para que seja transmitido para outras pessoas, entretanto, o que foi mudando ao longo do tempo foi o suporte de armazenamento que precisa estar em consonância com a realidade local, para a organização dos documentos e sua recuperação.

Os primeiros registros da informação foram encontrados nas cavernas, posteriormente, foram feitas escritas em tabletas de argila, onde além de ser um

lugar para registrar as informações já era possível certa mobilidade do suporte de armazenamento. Com o passar do tempo surgiram os rolos de papiro, que eram bem mais leves que os tablets de argila, e dando um “salto no tempo” chegamos ao papel que usamos até hoje, além das tecnologias recentes que proporcionam a elaboração de documentos digitais e a digitalização de documentos analógicos.

Já na Antiguidade as bibliotecas, possuíam um acervo híbrido, com vários tipos de documentos em suportes diferentes, nesta época elas já eram vistas como ambientes que deveriam reunir todo o saber, preconizando o desejo de se criar uma biblioteca universal, que pudesse armazenar todo o conteúdo produzido pelos humanos, entretanto, não era de fácil acesso a qualquer cidadão, representavam poder e *status* social a quem dela fosse curador ou pudesse acessar.

Com o passar dos séculos e principalmente com a invenção da imprensa, a produção informacional começou a crescer e a demanda por um espaço que pudesse acumular todo este volume de documentos ficou maior. Num primeiro momento a biblioteca atende a esta requisição básica, fato que se comprova quando nos reportamos ao significado do termo que surgiu na Grécia “cofre do livro”, que de acordo com Ribeiro (1996) servia tanto aos lugares onde os livros eram alocados como também às coleções de livros em si.

Obviamente que esta posição serve apenas como um pilar para captarmos a ideia do que seria uma biblioteca nos primórdios da sua existência e seu conceito envolve uma série de outras atribuições que excedem ao armazenamento ou mero “depósito de livros”.

Ribeiro (1996, p. 29) nos diz que,

[...] desde as origens, a função "serviço" - uso e pesquisa de informação - surge implícita no conceito de "biblioteca". Já na Antiguidade temos, portanto, perfeitamente conceptualizada uma realidade, que progressivamente se complexificou, tornando nítidas as suas componentes próprias, que permitem hoje afirmar, sem receio, que o conceito define um sistema de informação.

Segundo Milanesi (2002, p. 7) “a partir do momento em que o homem foi capaz de perceber que a produção e organização do produto a ser preservado são ações paralelas e simultâneas e de colocar em prática essa atividade de ordenamento estabeleceu a noção básica de biblioteca”; o que nos faz lembrar mais

uma vez, que a biblioteca é anterior à Idade Média, surgindo da necessidade de registrar, disseminar e recuperar informação.

3.1 A Biblioteca Universal

A primeira edição do dicionário Soares Amora (1997, p. 92) tem a seguinte definição para biblioteca “Coleção de livros, disposto de modo ordenado e em estantes especiais; nome dessas estantes ou do lugar onde se encontram”.

Com este conceito visualizamos que a descrição de biblioteca acima informada se assemelha ainda com o termo surgido na Grécia Antiga quando se iniciou o uso de tal nomenclatura. Utilizamos esta terminologia relativamente recente para começar a investigar as mutações sofridas pelas bibliotecas e entender como chegamos às bibliotecas digitais.

A procura através dos séculos pelo conhecimento universal reunido em um só lugar não é também uma novidade, desde a construção da famosa Biblioteca de Alexandria, provavelmente no século III a.C., buscava-se armazenar todo o conteúdo burocrático, rolos de papiro, imagens e ainda,

[...] parte do acervo desta biblioteca foi constituído principalmente a partir de um decreto de Ptolomeu III em que todos os navios que parassem em Alexandria tinham que entregar seus livros para serem copiados (após a cópia, às vezes devolviam-se os originais, às vezes as cópias). O objetivo era o de uma biblioteca que abrigasse a totalidade do conhecimento humano registrado.[...]. (ORTEGA, 2004, p. 23)

Mas como se sabe, apesar de ter resistido a vários ataques durante um grande período, ela acabou deixando de existir, ao contrário do sonho do ser humano em organizar e abrigar o grande volume de sua produção documental que só ascendeu com o transcorrer dos séculos. Nesta linha de raciocínio já no ano de 1895 estavam os advogados belgas, Paul Otlet e Henri La Fontaine que de acordo com Cunha (2008) criaram, em Bruxelas, o Instituto Internacional de Bibliografia e tinham o objetivo de “registrar em fichas a produção mundial de impressos: o Repertório Bibliográfico Universal, então inaugurado naquela cidade” (FONSECA, 2007, p. 4); para o qual o Brasil também contribuiu através do então diretor da Biblioteca Nacional, localizada no Rio de Janeiro, Manuel Cícero Peregrino da Silva que enviava dados da produção bibliográfica brasileira (FONSECA, 2007).

Esta iniciativa também durou por muito tempo, ainda assim, como destaca Sayão (2008, p. 4),

A ideia de um repositório que se desdobre ao infinito registrando e organizando todo o conhecimento humano parece ser um sonho obsessivamente renovado ao longo do tempo. [...] e as mais avançadas tecnologias de todas as épocas sempre estiveram a serviço da sua concretização.

Avançando um pouco mais no tempo, em 1945, nos deparamos com a proposta de Vannevar Bush, a então pretensiosa máquina *MEMEX*, segundo o próprio criador que a define como “um dispositivo em que o indivíduo armazenará seus livros, seus registros, suas anotações, suas comunicações. O dispositivo será mecanizado de modo a poder ser consultado com extrema velocidade e flexibilidade”. (BUSH, 1945, p. 101 *apud* SILVA; SÁ; FURTADO, 2005). Este invento se aproxima bastante do que é utilizado em bibliotecas digitais hoje, no ideal de que o MEMEX parece o que seria um computador atual equipado com acesso a internet e banco de dados no qual os aplicativos usariam hipertextos.

Desta forma pode-se solicitar qualquer livro de sua biblioteca e consultá-lo de forma mais fácil do que se tivesse de retirá-lo de uma estante. [...] Aparecerão formas completamente novas de enciclopédias feitas com uma trama de cadeias associativas prontas para serem introduzidas no MEMEX e ali ampliadas. (BUSH, 1945, p. 102-104).

Bush é considerado, por muitos estudiosos, o precursor das bibliotecas digitais, pois previu os tempos hodiernos, onde a velocidade da produção informacional é vertiginosa e gera a necessidade da existência de um dispositivo mecanizado que facilite o tratamento e administre o grande volume de informações de modo veloz e flexível, servindo de base para criações futuras; que também dariam continuidade às tentativas de construção de uma biblioteca universal, com acervos que vão além do concreto, real, sendo ele levado para uma dimensão onde possam ser depositados e recuperados sem estar necessariamente num ambiente físico e, ainda assim, possam estar disponíveis a todos com serviços prestados a distancia.

Para Pereira e Rutina (1999, p.13) “[...] o ideal da biblioteca universal, passa, é claro, pela cooperação e intercâmbio entre as bibliotecas”. Para elas o foco do desenvolvimento desta caminhada rumo à integração de documentos produzidos ao redor do mundo se concentra nas atividades de intercâmbio e cooperação entre as

bibliotecas, o que está diretamente relacionado ao fato das bibliotecas estarem ocupando o espaço virtual e cada vez mais integradas entre si através de redes de computadores.

Pereira e Rutina (1999, p. 14) associam o ideal da biblioteca universal a própria internet devido aos serviços prestados *on-line* e diversos banco de dados espalhados pela rede.

Felizmente, já desde a Guerra Fria, o mundo conta com uma poderosa ferramenta de acesso aos documentos e às informações, qual seja, a Internet, considerada uma grande biblioteca virtual. Isso, porque há milhares de bancos de dados, arquivos e serviços *on-line* independentes disponíveis aos usuários na rede mundial de computadores. Trata-se de uma rede muito especial, porque a sua utilização aproxima os países, ao mesmo tempo em que anuncia uma cultura global.

O advento da internet, sem dúvidas abriu um leque de oportunidades para as comunicações, rompendo com a linearidade da escrita e possibilitando associações semânticas; agregando-se às bibliotecas oportunizou e oportuniza a prosperidade do desejo das mentes criativas associadas ao estudo da informação no empreendimento da biblioteca universal; isto favoreceu também o desenvolvimento das bibliotecas digitais que puderam alcançar mais usuários, ligar seu banco de dados ao de outras bibliotecas em rede e passar a concentrar seus esforços em questões de interoperabilidade de dados e padronização, dentre outras ocupações que antes da *Web* estavam um tanto adormecidas e dificultavam o intercâmbio de informações.

Bush não conseguiu acompanhar toda a transformação que ele previa na primeira metade do século XX, mas suas ideias foram cruciais para o posterior desenvolvimento dos repositórios digitais e das “B.D.”¹⁷.

Podemos confirmar que estas últimas assim como as “*Webs*” vêm de uma série de transformações que começaram com o pensamento de indivíduos ao longo da história que almejavam concentrar toda produção intelectual escrita em um só espaço.

Pode-se perceber na literatura científica da área de Ciência da Informação e da Biblioteconomia, que há uma pluralidade conceitual sobre a definição de

¹⁷ B.D. são aqui consideradas Bibliotecas Digitais.

biblioteca digital, muitas vezes se comparando à virtual. Nesse sentido, é pertinente destacar que há diferenças nesses modelos propostos.

3.2 Biblioteca virtual ou digital?

A discussão sobre o virtual se faz necessária, pois muitas vezes há um problema conceitual verificado na literatura científica, sobre o tema biblioteca digital e biblioteca virtual, gerando dificuldades no seu entendimento, sobretudo, por bibliotecários.

Se considerarmos o virtual como algo intangível, avesso do que é concreto, poderemos seguir uma lógica onde a definição de virtual estaria intimamente ligada à oposição do real, já que este último está associado à tangibilidade e a expressão “realidade virtual” poderia nos causar um “ar paradoxal” e conflituoso. No entanto, não seguiremos este pensamento, pois segundo Pierre Lévy em diversas publicações tais como, “O que é virtual?” e “Cibercultura”, nos oferece subsídios para entendermos o virtual de outra forma.

De acordo com Lévy (1997, p. 8) a palavra virtual vem do latim medieval *virtualis*, e esta deriva de *virtus* que representa “força”, “potência”, em outras palavras, o virtual está muito mais ligado a uma força potencial, algo que potencialmente será, em exemplos práticos uma semente é potencialmente/virtualmente uma árvore e não o seu oposto. Nesse sentido, o autor coloca que o virtual faz contraposição com o atual e não com o real.

Já o virtual não se opõe ao real, mas sim ao atual. Contrariamente ao possível, estático e já constituído, o virtual é como o complexo problemático, o nó de tendências ou de forças que acompanha uma situação, um acontecimento, um objeto, ou uma entidade qualquer e que chama um processo de resolução: a atualização. Esse complexo problemático pertence à entidade considerada e constitui inclusive uma de suas dimensões maiores. O problema da semente, por exemplo, é fazer brotar uma árvore. (LEVY, 1995, p. 5)

Em consenso com tal afirmação podemos vincular o sentido de virtual com o digital, já que as informações em códigos binários (0 e 1) transferidas pelas redes de computadores, são representações do original as quais são ilegíveis pelo homem; enquanto dados, apesar de existirem realmente em determinado nó da rede, estão virtualmente presentes em diversos outros computadores que fizeram a requisição da mesma; estes últimos não tem acesso a informação de origem, apenas as

atualizações que se refletem em seus aparelhos eletrônicos quando solicitadas, mas que foi alterada de modo efetivo, somente em um ponto da rede onde encontra-se de fato armazenada.

[...] a digitalização da informação pode ser aproximada da virtualização. [...] A informação digital (traduzida para 0 e 1) também pode ser qualificada de virtual na medida em que é inacessível enquanto tal ao ser humano.[...] Um mundo virtual – considerado como um conjunto de códigos digitais – é um potencial de imagens, enquanto uma determinada cena, durante uma imersão no mundo virtual atualiza esse potencial em um contexto particular de uso. Essa dialética do potencial do cálculo e da exibição contextual caracteriza a maioria dos documentos ou conjuntos de informações de suporte digital. (LEVY, 1999, p.48-49)

Esta explanação sobre virtual e digital traz ao foco principal deste capítulo, as bibliotecas digitais, assunto este que gera polêmica, sobretudo, sob a ótica conceitual, pois as mesmas lidam com as informações digitais ou virtuais.

3.3 Afinal, o que é Biblioteca Digital?

Conhecemos um pouco do percurso que nos trouxe até as bibliotecas digitais, entretanto, é pertinente entendermos de fato o que é uma biblioteca digital, mas isso pode não ser uma tarefa tão simples, achamos um conceito único que nos responda a esta questão diante um tema ainda em discussão. Nesse sentido, apresentaremos algumas definições de bibliotecas digitais, segundo. Barker (1994), Marchiori (1997), Drabenstott e Burman, (1997), Lemos (1998), Costa (2000), Alvarenga (2001), Borges (2003), Toutain(2006), Cunha (2008) e por fim Sayão (2008) e (2009):

3.3.1 Conceitos atribuídos à Biblioteca Digital

A biblioteca digital segundo Barker (1994 *apud* MOREIRA, 1998, p. 28), tem como característica,

[...] uma coleção de documentos eminentemente digitais; independentemente se foram criados na forma digital ou digitalizados a partir de documentos impressos, permite, por meio do uso de redes de computadores, compartilhar a informação instantânea e facilmente.

Segundo Marchiori (1997, p. 5), a biblioteca digital se difere das demais, porque

a informação que ela contém existe apenas na forma digital, podendo residir em meios diferentes de armazenagem, como as memórias eletrônicas (discos magnéticos e óticos). Desta forma, a biblioteca digital não contém livros na forma convencional e a informação pode ser acessada, em locais específicos e remotamente, por meio de redes de computadores. A grande vantagem da informação digitalizada é que ela pode ser compartilhada instantânea e facilmente, com um custo relativamente baixo.

A biblioteca digital, biônica, transmitida eletronicamente em rede – a biblioteca do século XXI,

[...] vem a ser uma potente, transparente e universal provedora de informação, também chamada de biblioteca sem paredes. [...] possibilita acesso a um reservatório diversificado e infinito de dados e conhecimento.[...] Cerca vários tipos de recursos informacionais e os mais dinâmicos artefatos digitais. [...] Um grande número de usuários pode acessá-la simultaneamente por conexão a redes eletrônicas internacionais.[...] (DRABENSTOTT; BURMAN, 1997 *apud* COSTA 2000, p. 124).

Para Lemos (1998, p. 346), a biblioteca digital seria aquela que teria,

[...] além de seu catálogo, também os textos dos documentos de seu acervo armazenados de forma digital, permitindo sua leitura na tela do monitor ou sua importação (download) para o disco rígido do computador que funcione como porta de acesso à internet , sem desprezar toda a gama de opções que o sistema de hipertexto poderá oferecer em termos de interligação de sítios no universo da internet.

Costa (2000, p. 125), destaca o conceito de bibliotecas digitais bastante similar às bibliotecas convencionais/tradicionais, ao mencionar que as bibliotecas digitais possuem [...] uma herança das características da biblioteca tradicional baseada no suporte atômico da informação.[...].

Alvarenga (2001, p. 8) aponta dentre as definições de bibliotecas digitais, o destaque pelo conjunto de objetos digitais, e dessa forma, são construídos, a partir do “uso de instrumentos eletrônicos, concebidos com o objetivo de registrar e comunicar pensamentos, ideias, imagens e sons, disponíveis a um contingente ilimitado de pessoas, dispersas onde quer que a plataforma www alcance”.

A biblioteca digital surge-nos, assim, como uma meta-entidade capaz de aglomerar, não física mas virtualmente, os objectos de informação. Caracteriza-se por vários fenómenos entre os quais

destacamos o de desterritorialização e o de eliminação das categorias espaço-temporais. (BORGES, 2003, p. 663)

Toutain (2006, p. 16) destaca novamente a questão da base informacional das bibliotecas digitais, onde seus conteúdos são caracterizados por,

[...] texto completo em formatos digitais – livros, periódicos, teses, imagens, vídeos e outros -, que estão armazenados e disponíveis para o acesso, segundo processos padronizados, em servidores próprios ou distribuídos e acessados via rede de computadores em outras bibliotecas ou redes de bibliotecas da mesma natureza.

A biblioteca digital combina a estrutura e a coleta da informação tradicionalmente usadas por bibliotecas e arquivos, com o uso da representação digital tornada possível pela informática. “[...] À semelhança da biblioteca convencional, a biblioteca digital também inclui os princípios consagrados de como a informação é organizada”. (CUNHA, 2008, p.5)

Sayão (2009) aponta que além das diferentes visões relativas ao tempo em que está se conceitualizando as bibliotecas digitais, as diferenças também irão surgir de acordo com a área de atuação do profissional que lidar com ela.

Por envolver quesitos que atraem a atenção de arquivistas, bibliotecários e profissionais das tecnologias da informação e comunicação em geral, os horizontes de suas aplicabilidades se expandem ainda mais causando novas terminologias de acordo com o enfoque do profissional que dela faz uso. (SAYÃO, 2009, p.15)

Quando se trata da área de Biblioteconomia, de modo particular, as visões conceituais para bibliotecas digitais recaem sobre as questões de organização e tratamento dos recursos informacionais, ações basilares em bibliotecas convencionais/tradicionais.

A comunidade de Biblioteconomia dará ênfase ao lado organizacional das bibliotecas digitais fazendo destas uma extensão dos serviços disponibilizados pelas bibliotecas tradicionais, uma nova infraestrutura tecnológica para potencializar a disseminação de informação e conhecimento e não como máquina ou aparato tecnológico, diferente da visão dos cientistas da computação que a enxergam como uma extensão de aplicativos em rede, que à medida que se especificam as áreas da computação a visão também muda, ou seja, para o profissional que trabalhar com banco de dados a verá como uma continuação deste ramo, já os que trabalham com multimídia a verá desta forma e assim sucessivamente. (SAYÃO, 2008, p.12).

Para essa pesquisa utilizaremos a definição de biblioteca digital empregada por Sayão (2009), pois acreditamos que essa se adeque aos propósitos deste trabalho e para o encaminhamento e construção da *Web Semântica*.

A ideia de biblioteca digital tem muitas faces, mas nenhuma delas a define completamente e esgota todos seus significados. As definições de biblioteca digital se reconfiguram de acordo com os seus números protagonistas que se espalham por muitas áreas. [...] O breve passado das bibliotecas digitais não foi capaz de resolver essas ambigüidades, quem sabe o futuro seja rápido em harmonizá-las. (SAYÃO, 2009, p. 15-16).

O que se pode concluir com este breve revisitar de conceitos na literatura científica da Ciência da Informação e da Biblioteconomia é que ainda hoje há uma dificuldade em encontrar apenas um termo que envolva a grandiosidade de aplicações referentes às bibliotecas digitais e, por mais que se crie um conceito, ele não englobará toda a capacidade informacional que está contida neste tipo de ambiente informacional.

3.3.2 Metodologia para a construção de Bibliotecas Digitais

A biblioteca digital é composta por documentos multimídia que provavelmente estarão disponíveis através de uma rede de computadores, para que eles sejam recuperados da melhor maneira pelos usuários, ou seja, com o máximo de precisão, com base no estudo feito com os diversos autores já citados ao longo deste capítulo formulamos as seguintes indicações para a construção de uma biblioteca digital:

1. deverão estar bem detalhados através de uma representação descritiva, fazendo uso de metadados e linguagens de marcação;
2. definir um banco de dados confiável, analisando sempre a necessidade da instituição na qual a biblioteca está vinculada e se é preciso utilizar soluções gratuitas ou pagas;
3. indexar uma grande quantidade de palavras-chave que proporcionem a rápida identificação do documento;
4. a escolha de um formato que facilite a padronização dos dados favorecendo o intercâmbio;
5. estar conectada com a internet;

6. pensar em uma interface de fácil utilização, que possa atrair ao invés de repelir usuários;
7. ter computadores;
8. definir o perfil do seu usuário;
9. obter um acervo que atenda às expectativas do público-alvo;
10. ter profissionais capacitados para lidar tecnologias da informação atuais;
11. divulgar a biblioteca digital.

Os passos acima relacionados são apenas alguns itens que precisam ser questionados quando se pensa em construir uma biblioteca digital, nem todos são obrigatórios, mas representam elementos para a boa prática na elaboração e também não estão limitados a estes. Outros autores relacionam listas de características atribuídas às bibliotecas digitais, que podem ser observadas para a sua concepção, a exemplo de Cunha (1999) que destaca:

- a) acesso remoto pelo usuário, por meio de um computador conectado a uma rede;
- b) utilização simultânea do mesmo documento por duas ou mais pessoas;
- c) inclusão de produtos e serviços de uma biblioteca ou centro de informação;
- d) existência de coleções de documentos correntes onde se pode acessar não-somente a referência bibliográfica, mas também o seu texto completo. O percentual de documentos retrospectivos tenderá a aumentar à medida que novos textos forem sendo digitalizados pelos diversos projetos em andamento;
- e) provisão de acesso em linha a outras fontes externas de informação (bibliotecas, museus, bancos de dados, instituições públicas e privadas);
- f) utilização de maneira que a biblioteca local não necessite ser proprietária do documento solicitado pelo usuário;
- g) utilização de diversos suportes de registro da informação tais como texto, som, imagem e números;

h) existência de unidade de gerenciamento do conhecimento, que inclui sistema inteligente ou especialista para ajudar na recuperação de informação mais relevante.

Em outras palavras, para construir uma biblioteca digital precisa-se levar em consideração todo o estudo que deve ser realizado para as bibliotecas tradicionais observando as especificidades do meio digital, os equipamentos e infraestrutura tecnológica para o bom funcionamento do sistema, além do conhecimento necessário em tecnologia por parte do profissional bibliotecário que ficará responsável pela gestão desta biblioteca.

Nesse sentido o quadro 4 aponta as principais habilidades requeridas ao profissional da informação para lidar com as Bibliotecas digitais, conforme abordadas por Choi e Rasmussen (2006) apud Cunha (2008).

Quadro 3 – Habilidades profissionais necessárias em Bibliotecas Digitais.

Tecnologia	Biblioteconomia	Outros
Arquitetura e software de biblioteca digital (4,52)	Necessidade dos usuários (4,52)	Comunicação interpessoal (4,60)
Normas e padrões de qualidade (4,33)	Preservação digital (4,42)	Gerenciamento de projetos (4,56)
Linguagens de marcação (4,04)	Catálogo, metadados (4,0)	Problemas legais (3,90)
Sistema de gerenciamento de banco de dados (3,77)	Indexação (3,71)	Captação de recursos (3,88)
Web design (3,71)	Desenvolvimento de coleções (3,71)	Ensino e aprendizagem para grupos (3,52)

Fonte: CHOI; RASMUSSEN (2006) apud Cunha (2008).

Nota: entre parênteses está a média na escala Likert de cinco pontos.

Nesse quadro Choi e Rasmussen (2006) apud Cunha (2008) apresentaram alguns profissionais de biblioteconomia questões do conhecimento que seriam importantes para o bibliotecário que pretende desenvolver uma biblioteca digital.

Para essa análise foi utilizada a escala de Likert, que consiste na pontuação máxima de 5 pontos e representa o maior grau de relevância, obtendo os referidos resultados entre parênteses.

Com tudo isto, ratificamos a importância das discussões acerca das bibliotecas digitais, que independente de agregar um conceito mais restrito ou mais abrangente, exige do profissional maior capacitação tecnológica para que possa desenvolver práticas eficientes que atendam as requisições dos usuários e, de preferência, supere as suas expectativas, colaborando assim com a evolução da práxis bibliotecária e o conhecimento organizacional deste profissional da informação e possa contribuir com a ordenação informacional no ambiente *Web*.

Neste sentido, o capítulo seguinte apresenta o estudo de algumas práticas atribuídas ao desenvolvimento baseado na *Web Semântica*, em suas camadas básicas dando enfoque a representação descritiva dos documentos, atividade realizada nas bibliotecas digitais, sob a ótica dessa nova forma de composição para a *Web* podendo os bibliotecários usufruir dos instrumentos tecnológicos que os auxiliem na missão de atender a mais usuários e da melhor forma possível.

4 AS TECNOLOGIAS DAS CAMADAS-BASE DA WEB SEMÂNTICA E SUAS CORRELAÇÕES COM AS BIBLIOTECAS DIGITAIS

Nos primeiros capítulos vimos o caminho evolutivo da *Web*, no sentido de se consolidar e aumentar o uso da comunicação via rede de computadores através do apelo social e de instrumentos tecnológicos que trabalhassem em função disto. A atração de um grande quantitativo de usuários e da produção e circulação de documentos no ambiente virtual, paralelamente a estes acontecimentos foi observada a inserção das bibliotecas no espaço digital tentando acompanhar as tecnologias vigentes, se inserindo de forma gradativa, mas sólida, talvez rumo ao que um dia nomearam de biblioteca universal.

Nesta linha de raciocínio continuamos no progresso da *Web*, das bibliotecas digitais e dos benefícios desta parceria em função do avanço organizacional na troca de documentos digitais via telemática, chegando então, na terceira geração da *Web*, conhecida como *Web Semântica* ou *Web 3.0*, trazendo o foco desta pesquisa que tem por objeto o estudo de instrumentos concentrados nas camadas Unicode/URI, XML e RDF desta *Web*, visando uma congruência de tecnologias com as bibliotecas digitais.

No intuito de termos uma base substancial para compreendermos as especificidades das camadas em questão nas discussões futuras, trabalharemos inicialmente no entendimento sob uma visão *top-down*, ou seja, a *Web Semântica* do geral para o específico, supondo deste modo, ser a melhor forma de assimilação. Precedendo, ainda tal ótica, estará uma sucinta disposição sobre a origem e as discussões terminológicas acerca da *Web Semântica*.

4.1 Acerca da *Web Semântica*

A palavra semântica deriva do grego *Semantiká*, (plural neutro de *Semantikós*, que por sua vez deriva de *Sema*, sinal), tendo como objetivo o estudo do significado, um nome que envolve bem as situações-problema que a *Web 3.0* pretende resolver.

A *Web Semântica* foi uma proposta que emergiu efetivamente para toda a comunidade científica no artigo “*The Semantic Web: a new form of Web content that*

is meaningful to computers will unleash a revolution of new possibilities”,¹⁸ publicado na revista *Scientific American*, em 2001, onde os autores Tim Berners-Lee, Ora Lassila e James Handler anteviam situações em que num futuro próximo existiria uma espécie de *Web* integrada onde os agentes de *software* analisariam o conteúdo das páginas nas quais realizariam suas buscas baseados na descrição bem detalhada contida nelas, realizando de forma automática refinamentos na pesquisa pedida pelo usuário resultando numa busca semântica, com respostas mais precisas, evitando o trabalho manual feito pelo humanos.

Em termos práticos, por exemplo, uma pessoa está com dor de dente e precisa ir ao dentista, recorre à *Web* para pesquisar um consultório mais próximo da sua casa, que tenha a melhor conceituação perante os clientes e que os dentistas atendam aos sábados pela manhã. O retorno desta busca não seria preciso na *Web* atual, (conhecida como *Web* Sintática, *Web* Social ou *Web* 2.0), haja vista que ela está baseada em palavras-chave que podem ser compreendidas apenas por seres humanos e não por máquinas; os agentes de *software* não conseguem “entender” o significado das palavras nem atribuir sentido a elas, para que deste modo pudessem relacionar os dados contidos em outras páginas podendo assim escolher, através do cruzamento de informações semânticas, as que melhor atendessem a solicitação requerida, ao invés de somente retornar um grande volume de páginas que mencionam pelo menos uma das palavras contidas na solicitação acima.

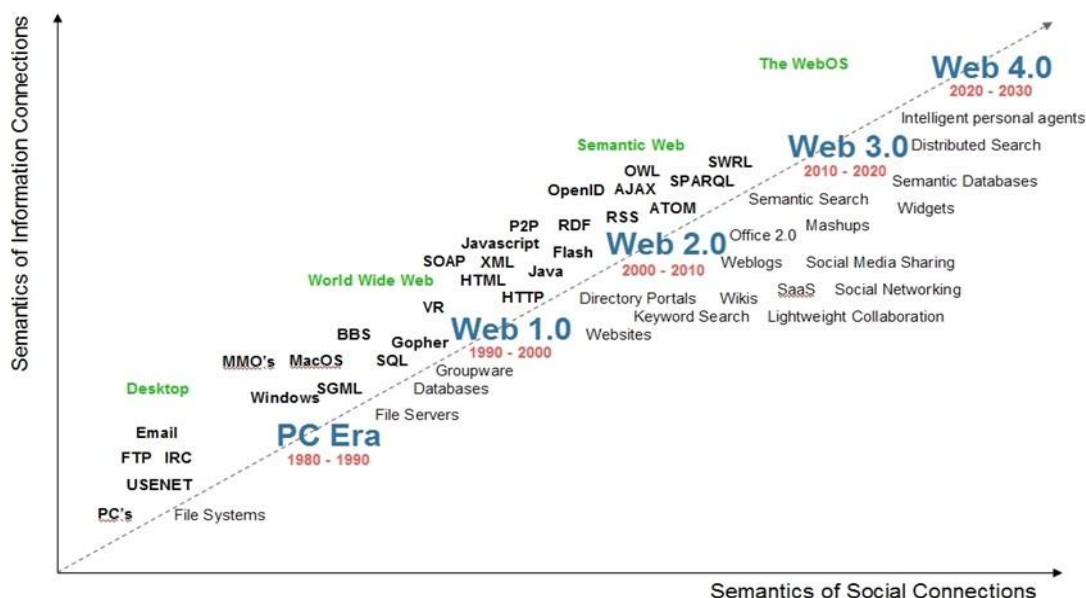
A *Web* Semântica não é uma *Web* separada, mas uma extensão da atual em que a informação é dada com significado bem definido, permitindo que computadores e pessoas trabalhem melhor em cooperação. Os primeiros passos para tecer a *Web* Semântica a partir da estrutura da *Web* existente já estão em andamento. Em um futuro próximo, este desenvolvimento vai inaugurar uma nova e significativa funcionalidade para as máquinas se tornarem muito mais capazes de processar e “entender” os dados que elas simplesmente mostram no presente. (BERNERS-LEE; LASSILA; HENDLER, 2001, p.2).

A *Web* Semântica é uma ampliação da atual 2.0 que por sua vez, deriva da primeira versão (conforme explicitado no capítulo 2), iniciada a partir das transmissões via FTP – *File Transfer Protocol* (ver figura 4). Quanto a ser considerada como *Web* 3.0 ainda gera controvérsias entre teóricos das áreas relacionadas às Tecnologias da Informação e Comunicação (TIC). No caso de

¹⁸ Tradução: “*Web* Semântica: um novo formato de conteúdo para a *Web* que tem significado para computadores que vai desencadear uma revolução de novas possibilidades”.

Santos e Alves (2009), elas mostram através do quadro de Méndez Rodríguez (2007) que a *Web Semântica* está na verdade intermediária entre as *Webs 2.0* e 3.0, para dar uma espécie de estruturação aos dados e suporte tecnológico para a aplicação de outros instrumentos e criação de novos ambientes informacionais propostos na *Web 3.0*.

Figura 5 – Evolução da Web 1.0 a 4.0.



Fonte: Méndez Rodríguez (2007) *apud* Santos e Alves (2009).

Analisando a figura 4, nota-se que a *Web Semântica* é considerada como um conjunto de recursos tecnológicos, linguagens-padrão para ontologias, metadados etc., que complementa as tecnologias já existentes na *Web 2.0*, tais como linguagem XML, *framework* RDF, dentre outras que podem ser usadas para o tratamento semântico da informação, com o objetivo de ser uma estrutura-suporte para a *Web 3.0*.

[...] pode-se considerar que a Web 3.0 não se constitui como a Web Semântica, sendo esta a denominação dada para a estrutura tecnológica criada para estabelecer maior nível semântico aos dados; seu funcionamento envolve a implantação de ferramentas tecnológicas e métodos de representação da informação para o estabelecimento de raciocínio sobre os dados e, a partir daí, facilitar a recuperação e a construção de ambientes inteligentes. Já a Web 3.0 se constitui como uma denominação para um período de

evolução da Web marcado pela criação de ambientes informacionais altamente especializados e que só funcionarão efetivamente a partir da implantação da estrutura da Web Semântica.[...] Sendo assim, percebe-se que a Web 3.0 só poderá se efetivar a partir da implantação de uma estrutura tecnológica anterior, proporcionada pela Web Semântica [...]. (SANTOS; ALVES, 2009, p. 2).

Por outro lado, se o entendimento evolutivo da *Web* seguir outra linha de pensamento, que pode ser exemplificada na mesma figura 4, encontraremos subsídios para associação dos termos *Web Semântica* e *Web 3.0*. Vejamos a passagem da primeira geração da *Web* para a segunda, houve grande transformação, já que a primeira era apenas o início desta grande rede e seu foco era apenas se consolidar diante dos ainda tímidos (em relação aos padrões atuais) usuários da internet; diferente da 2.0, que agregou outras tantas tecnologias a exemplo do uso crescente de linguagens multiplataformas como a JAVA, recursos tecnológicos como os *plug-ins*, Really Style Sheets (RSS) e *Mashups*, navegadores mais rápidos, criação de redes sociais e muitas outras aplicações, que foram criadas a partir da base de hipertextos já existente na 1.0.

Pensando por este ângulo a *Web 2.0* “preparou” a *Web* para atrair um grande número de usuários e para a interação entre eles, mas de forma um tanto anárquica quanto à organização da informação. Deste modo, não poderia a *Web 2.0* ser considerada uma *Web Sintática*, que prepara o ambiente para nova geração da *Web*, a real *Web 2.0*? Ou ao contrário, já que a *Web Semântica* tem uma nova proposta para a organização da *Web* atual não poderia assim ser considerada *Web 3.0*?

Aghaei, Nematbakhsh e Farsani (2012, p. 5, tradução nossa) preferem associar os dois termos em questão,

[...] Web 3.0 tenta vincular, integrar e analisar dados de vários dados, definindo-os para obter um fluxo de novas informações, é capaz de melhorar a gestão de dados, suporte a acessibilidade da internet móvel, simular criatividade e inovação, incentivar o fenômeno da globalização, aumentar a satisfação dos clientes e ajudar a organizar a colaboração na web social. Web 3.0 é também conhecida como Web Semântica. A Web semântica foi pensada por Tim Berners-Lee, inventor da World Wide Web. Há uma equipe dedicada ao consórcio World Wide Web (W3C), trabalhando para melhorar, ampliar e padronizar o sistema, idiomas, publicações e ferramentas que já foram desenvolvidas. Web semântica é uma rede que pode demonstrar as coisas numa abordagem que o computador possa entender. O principal objetivo importante da Web Semântica é tornar a Web legível por máquinas, e não por seres humanos.

A partir de tal impasse entendemos que a discussão sobre a terminologia a ser adotada ainda gera controvérsias, pois o termo *Web 3.0*, segundo os autores supracitados, surgiu em 2006 com Jonh Makoff, do *The New York Times*, para indicar a terceira geração da *Web*, que tem por objetivo “definir a estrutura de dados e ligá-los de forma mais eficaz para a descoberta, automação, integração e reutilização em diversas aplicações” (AGHAEI; NEMATBAKHSH; FARSANI, 2012, p. 5); conceito que se assemelha às propostas da terminologia “*Web Semântica*”, termo que se faz mais notável em 2001, com o artigo de Berners-Lee.

Em virtude de recentes artigos publicados na área das TIC com a ótica semelhante às de Aghaei, Nematbakhsh e Farsani (2012), assumimos a postura de associar os termos *Web 3.0* à *Web Semântica*, por também entendermos serem sinônimos. Ainda assim, por hora, não dá para colocar esta visão como a principal, pois, isto requer maior análise das terminologias e este não é o foco desta pesquisa.

Superando as questões terminológicas nos ateremos agora em estudar a *Web Semântica* em si. Segundo Bulcão Neto, Prazeres e Pimentel (2006), para galgar seus objetivos ela se baseia em três elementos fundamentais, tais como os padrões de metadados, ontologias e especificações de representação sintática, estrutural, lógica e semântica de informações. Objetivos estes, que estão centrados no uso de instrumentos tecnológicas, distribuídas nas camadas específicas da sua estrutura.

Nesta *Web* “inteligente”, dados descreverão dados de modo organizado e de fácil acesso, mas “introduzir semântica à *Web* não é o mais difícil; o problema maior será criar padrões que sejam usados globalmente.” (GOÑI; FERNANDES; LUCENA, 2002, p. 5), para que possa haver o intercâmbio de informações entre as bases de dados. Esta dificuldade também é comentada por Breitman (2005, p. 5):

No entanto, acredita-se que o maior fator de sucesso e do crescimento da internet seja a liberdade que fornece a seus usuários. Em um mesmo ambiente temos *sites* sofisticados, construídos com a ajuda de especialistas, cuja informação serve como referência para empresas de todo o país, coabitando com páginas pessoais, construídas por leigos e que contém informações triviais sobre seus autores, familiares e bichos de estimação. É muito difícil imaginar que nesse ambiente quase anárquico possa prevalecer um único modelo de organização.

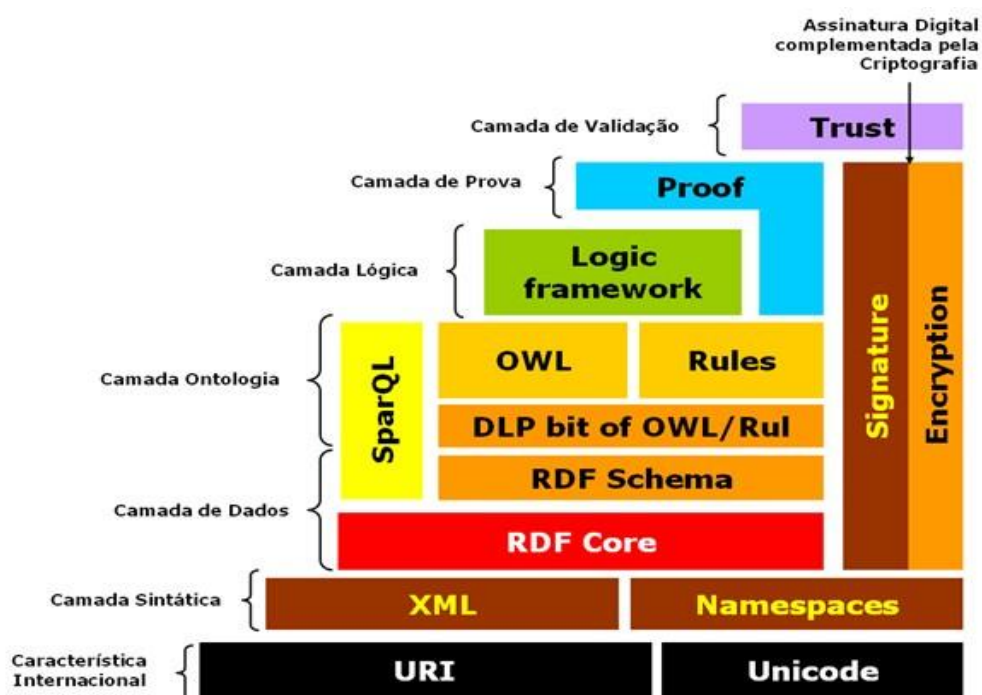
Dessa forma, é preciso um esforço coletivo para definir normas e padrões de organização no meio virtual, visando chegar-se aos padrões de intercâmbio de

dados propostos e, conforme citado pelos autores acima, não será algo fácil. Para tanto, é necessário o conhecimento da arquitetura e de conceitos prévios em tecnologias da informação para o manuseio dos instrumentos computacionais e informacionais. Neste momento, conheceremos um pouco da estrutura da *Web Semântica*.

4.2 Estrutura da *Web Semântica*

Para melhor compreendermos a *Web Semântica*, a figura 5 apresenta o delineamento da estrutura em camadas e os mecanismos tecnológicos que o sustenta.

Figura 6: Arquitetura em camadas da *Web Semântica*



Fonte: Castro (2012, p. 48)

A figura 6 representa a adaptação da estrutura em camadas proposta por Berners-Lee em 2005, para que se desenvolvam aplicações baseadas em semântica, com o foco na organização dos dados, uma espécie de “passo a passo” para construção de um ambiente interligado com metadados padronizados e ontologias que atribuem significado aos dados contidos na *Web*. Observa-se, com

isto, que é preciso um esforço conjunto entre as tecnologias contidas nas camadas sobrepostas, para que haja o estabelecimento efetivo da *Web Semântica*.

O enfoque desta pesquisa está concentrado nas tecnologias utilizadas nas camadas baixas, apontadas por Castro (2012), que são característica internacional, camada sintática e camada de dados. Estas camadas preparam a representação descritiva dos dados facilitando a atribuição das camadas subsequentes, simbolizando a base da estrutura, o que reflete na relevância do estudo destas.

Para termos uma visão geral sobre a atuação de cada camada Castro (2012, p. 48-50) nos traz as seguintes explicações:

- **Característica Internacional:** composta por padrões de descrição e identificadores universais, tais como a URI (*Uniform Resource Identifier*) e UNICODE, que são responsáveis por uma identificação e localização mínima do recurso informacional na rede;
- **Camada Sintática:** responsável pelo estabelecimento da sintaxe de descrição dos dados, essa camada é composta pela linguagem XML, *namespaces* e também pelo XML Schema, que permite uma melhor estruturação dos dados e consequentemente uma melhor visualização do conteúdo dos recursos pelos agentes de software;
- **Camada de Dados:** essa camada trabalha com a representação, o processamento e a codificação dos metadados. É composta pela arquitetura de metadados e/ou modelo de dados RDF (RDF Core) – Resource Description Framework e o RDF Schema, que são ferramentas responsáveis por expressar significados e promover a interoperabilidade. É importante destacar que nesta camada e na camada anterior está implícito o uso de padrões de metadados para garantir a representação dos recursos informacionais;
- **Camada de Ontologia:** o uso de ontologias permite o estabelecimento da semântica, ou o significado dos dados descritos e representados pelos metadados. A definição dos conceitos nessa camada servirá para os agentes inteligentes “compreenderem” a semântica dos dados, e a linguagem recomendada pelo World Wide Web Consortium (W3C) é a OWL – Ontology Web Language. Vale destacar que a SparQL é uma linguagem computacional utilizada para realizar consultas sobre estruturas RDF, incluindo estruturas descritas por meio da linguagem OWL. Como não possui mecanismos de inferência, possibilita manipular apenas informações armazenadas em um meio físico qualquer, em outras palavras, define as relações entre os dados armazenados na camada de estrutura.
- **Camada Lógica:** é responsável por uma recuperação mais eficiente devido ao uso de agentes, regras (presentes também na camada de ontologia) e mecanismos de inferência, mas para isso contam como base as camadas responsáveis pela estruturação, representação e

estabelecimento semântico dos dados, para posteriormente relacionar e processar as informações de forma lógica. Além disso, define regras lógicas que possam ser verificadas computacionalmente, permitindo a realização de inferências automáticas e a verificação do nível de coerência lógica dos recursos.

- **Camada de Prova:** responsável pelo intercâmbio entre agentes; é nessa camada onde serão processadas as definições lógicas pelos agentes para a construção da prova, que será criada para verificar se a dedução do agente estava correta;
- **Camada de Validação:** essa última camada é responsável pelo estabelecimento de verdades, autenticidade, confiabilidade e validade dos dados recuperados, garantindo aos agentes que raciocinam sobre os dados a veracidade da informação.

Dentro de cada camada existe um conjunto de outras tecnologias que darão suporte às camadas seguintes, daremos agora mais atenção as três primeiras camadas:

4.2.1 Camada Internacional (Unicode/ URI)

Esta camada mostra-se fundamental para o início do desenvolvimento semântico dos dados na *Web*, pois nela estão concentrados os primeiros esforços para a elaboração nesta arquitetura, com o fundamento de estabelecer um padrão universal para os caracteres para que possam atingir a interoperabilidade entre sistemas diversos e associar a informação a um único endereço, atribuindo assim, uma identidade onde a informação possa ser facilmente localizada, como conjunto de caracteres padrão da *WWW*. O *World Wide Web Consortium (W3C)* estabeleceu a utilização do *UNICODE*, antes mesmo das pesquisas em *Web Semântica*, observemos a seguir este padrão em detalhes.

4.2.1.1 O Padrão Unicode

Atualmente coordenado pelo *Unicode Consortium*, uma organização sem fins lucrativos, o *Unicode* é um padrão de caracteres que visa unificar os sistemas de representação de várias línguas em uma única normatização tentando eliminar as discrepâncias encontradas na comunicação entre sistemas de informação

automáticos que utilizam diferentes códigos de caracteres, o que contribui para dificultar o intercâmbio de dados.

Os computadores basicamente só compreendem números, nós só fazemos uso de caracteres por eles estarem associados a uma série numérica que corresponde às letras ou caracteres especiais (acentos gráficos, parênteses, símbolos matemáticos etc.) em diversos sistemas de numeração. A exemplo do binário, octogonal e hexagonal mais utilizados na informática, em especial o primeiro, e o mais comum no dia-a-dia, o decimal.

A codificação da tabela ASCII representada na figura 6, que tem por objetivo identificar caracteres através de uma associação com determinada numeração, tal como o Unicode, porém de modo muito mais restrito, pois, este padrão não engloba a quantidade e heterogeneidade dos glifos¹⁹ existentes em cada língua humana; nenhum sistema de codificação já o conseguiu, mas o Unicode tenta executá-lo com maior vantagem que o ASCII. (UNICODE, Inc. 2008).

O padrão Unicode abrange também a tabela ASCII, onde cada caractere é representado por um *byte*, um conjunto de zeros e uns (0 e 1), que suporta números binários de 0 a 255, ou seja, a cada caractere é atribuído um número entre 0 e 255, tendo em vista que sistemas baseados em ASCII suportam apenas 256 caracteres, um número muito aquém da quantia requerida para abarcar os glifos das diversas línguas existentes no mundo.

O Unicode consegue estender o ASCII, codificando a maioria dos caracteres mundiais, amplificando a representação de caracteres para a *range*,²⁰ de 0 a 10FFF hexadecimal. Esse número será expresso em uma das implementações num dos Formatos de Transformação Unicode (UTF - *Unicode Transformation Format*) que será aplicado adequando-se ao dispositivo em particular.

Dentre os Formatos de Transformação Unicode existem o UTF-8, UTF-16 e UTF-32, o primeiro é um formato que varia sua largura, podendo conter de um a quatro bytes, onde estes consistem de oito bits²¹ em sequências de um, dois, três ou quatro bytes, a depender do caractere a ser codificado, este modelo é amplamente adaptável a sistemas baseados em ASCII e devido a poucas diretrizes que façam

¹⁹ Glifo: em tipografia, é uma figura que dá um tipo de característica particular a um símbolo específico. Um glifo é um elemento da escrita.

²⁰ Range: neste caso, é entendido como campo de atuação, o intervalo entre 0 e 10FFF.

²¹ Bits: significa Dígito Binário ou *Binary Digit*, em inglês.

necessário o uso dos padrões subsequentes, principalmente do UTF-32. O UTF-8 é uma codificação bastante utilizada na internet, sistemas UNIX e banco de dados.

Figura 7 – Códigos da Tabela ASCII

Binário	Decimal	Hexa	Glifo	Binário	Decimal	Hexa	Glifo	Binário	Decimal	Hexa	Glifo
0010 0000	32	20		0100 0000	64	40	@	0110 0000	96	60	`
0010 0001	33	21	!	0100 0001	65	41	A	0110 0001	97	61	a
0010 0010	34	22	"	0100 0010	66	42	B	0110 0010	98	62	b
0010 0011	35	23	#	0100 0011	67	43	C	0110 0011	99	63	c
0010 0100	36	24	\$	0100 0100	68	44	D	0110 0100	100	64	d
0010 0101	37	25	%	0100 0101	69	45	E	0110 0101	101	65	e
0010 0110	38	26	&	0100 0110	70	46	F	0110 0110	102	66	f
0010 0111	39	27	'	0100 0111	71	47	G	0110 0111	103	67	g
0010 1000	40	28	(0100 1000	72	48	H	0110 1000	104	68	h
0010 1001	41	29)	0100 1001	73	49	I	0110 1001	105	69	i
0010 1010	42	2A	*	0100 1010	74	4A	J	0110 1010	106	6A	j
0010 1011	43	2B	+	0100 1011	75	4B	K	0110 1011	107	6B	k
0010 1100	44	2C	,	0100 1100	76	4C	L	0110 1100	108	6C	l
0010 1101	45	2D	-	0100 1101	77	4D	M	0110 1101	109	6D	m
0010 1110	46	2E	.	0100 1110	78	4E	N	0110 1110	110	6E	n
0010 1111	47	2F	/	0100 1111	79	4F	O	0110 1111	111	6F	o
0011 0000	48	30	0	0101 0000	80	50	P	0111 0000	112	70	p
0011 0001	49	31	1	0101 0001	81	51	Q	0111 0001	113	71	q
0011 0010	50	32	2	0101 0010	82	52	R	0111 0010	114	72	r
0011 0011	51	33	3	0101 0011	83	53	S	0111 0011	115	73	s
0011 0100	52	34	4	0101 0100	84	54	T	0111 0100	116	74	t
0011 0101	53	35	5	0101 0101	85	55	U	0111 0101	117	75	u
0011 0110	54	36	6	0101 0110	86	56	V	0111 0110	118	76	v
0011 0111	55	37	7	0101 0111	87	57	W	0111 0111	119	77	w
0011 1000	56	38	8	0101 1000	88	58	X	0111 1000	120	78	x
0011 1001	57	39	9	0101 1001	89	59	Y	0111 1001	121	79	y
0011 1010	58	3A	:	0101 1010	90	5A	Z	0111 1010	122	7A	z
0011 1011	59	3B	;	0101 1011	91	5B	[0111 1011	123	7B	{
0011 1100	60	3C	<	0101 1100	92	5C	\	0111 1100	124	7C	
0011 1101	61	3D	=	0101 1101	93	5D]	0111 1101	125	7D	}
0011 1110	62	3E	>	0101 1110	94	5E	^	0111 1110	126	7E	~
0011 1111	63	3F	?	0101 1111	95	5F	_				

Fonte: <http://www.aprender eletrônica.com>

O UTF- 16 segue a mesma linha de raciocínio, porém sua largura varia em unidades de 16 bits, preferida em aplicativos Microsoft Windows e em casos mais excepcionais, onde haja a necessidade do uso de caracteres raros implementa-se em UTF-32, aumentando a escala de unidades de dados a serem utilizados para 32

bits, exemplo comparativo entre os três tipos de UTF, conforme figura 7. Algumas vantagens relativas ao uso do padrão Unicode em banco de dados é descrita da seguinte forma no site da Microsoft, acessado no ano de 2013:

- Se os dados de origem e destino forem tipos de dados Unicode, o uso do formato de caractere Unicode preservará todos os dados de caractere.
- Se os dados de origem e destino não forem tipos de dados Unicode, o uso do formato de caractere Unicode minimizará a perda de caracteres estendidos nos dados de origem que não poderão ser representados no destino.

Figura 8 – Comparativo entre os Formatos de Transformação Unicode

CARACTERE	UTF-8	UTF-16	UTF-32
Letra Maiúscula Latina “A”	0x41	0x0041	0x00000041
Letra Grega Maiúscula “ALFA”	0xCD 0x91	0x0391	0x00000391
Letra Itálica Antiga	0xF0 0x80 0x83 0x80	0xDC00 0xDF00	0x00010300

Fonte: Elaborado pela autora.

Contribuindo para a uniformização dos caracteres na rede, o Unicode é uma peça-chave para o início do desenvolvimento sob o prisma da *Web Semântica*. Segundo Neto (2006 *apud* FRANCISCATO, 2010, p. 32),

O Unicode é fundamental para a *Web Semântica*, pois seus níveis superiores não precisam se preocupar em decodificar dados e transformá-los de um formato de representação para outro antes de enviar ou receber mensagens. Toda a informação enviada pela *Web Semântica* deve estar com seus respectivos identificadores numéricos definidos pelo Unicode. Portanto o Unicode pode ser transferido de uma plataforma para outra, pois não importa a forma

como ele foi armazenado em uma máquina, mas sim, o seu significado, que é padrão, tornando-o independente também do Sistema Operacional.

Observamos a relevância do uso do padrão Unicode nesta etapa inicial da arquitetura, mas não podemos esquecer que a camada em foco neste momento, agrega outra tecnologia para que, unidas, colaborem para o favorecimento da internacionalização e identificação dos caracteres e dados, a chamada URI (*Uniform Resource Identifiers*), como veremos mais sobre ela na subseção seguinte.

4.2.1.2 Uniform Resource Identifiers – URI

Um Identificador Uniforme de Recursos ou URI também tem o uso recomendado para o desenvolvimento de aplicativos com enfoque semântico. Conforme a arquitetura apresentada por Berners-Lee em 2005, mostrada com a adaptação de Castro (2012) na figura 5, ele está na camada mais “primitiva” da *Web Semântica* trabalhando em parceria com o Unicode para a inicialização da identificação, interoperabilidade e localização dos dados, evitando a incompatibilidade entre padrões de caracteres e identificando a localização de recursos na rede de computadores.

Especificamente o URI tem como objetivo encontrar o endereço de recursos informacionais digitais, ele referencia um recurso de modo único e global, subdivide-se em *Uniform Resource Locator* - Localizador Uniforme de Recursos (URL) e Nome Uniforme de Recursos (URN). Na sua forma mais conhecida, a URL, identificam-se recursos através da rede em parceria com o protocolos como DNS – *Domain Name System*, um sistema de gerenciamento de nomes hierárquico, que entre outros, associa um nome único à URL e HTTP - protocolo de aplicação responsável pelo tratamento de pedidos e respostas entre cliente e servidor na *World Wide Web* . Um exemplo de URL é o endereço *Web* da Fundação Biblioteca Nacional: <http://www.bn.br/portal/>, que descreve o local onde está armazenado geograficamente (.br) e o caminho da página a ser exibida no momento; neste caso, /portal representa a página principal. Com isto objetiva-se localizar o recurso mas não encontrar objetos ou nomes dentro do mesmo .

Quanto ao URN, Castello Branco Neto (2006, p. 27), tem a seguinte consideração:

[...] utilizado para definir nomes únicos e persistentes para recursos em um *namespace*, independente de onde ele esteja armazenado. De forma semelhante a uma URL, um URN pode se referir a um elemento acessível via *Web*, como documentos eletrônicos, mas pode, também, referenciar objetos ou pessoas que não estão acessíveis pela rede ou, ainda, conceitos abstratos como “Autor” e “Ano de Publicação” de uma obra.

Neste prisma o mesmo autor afirma que na *Web Semântica* há uma inclinação para que a URN tenha o mesmo *status* da URL em função da importância de ir além da identificação do local, onde a informação esteja armazenada, mas que se saiba do que ela está tratando, ou seja, do seu significado. Berners-Lee (2005, p.7) alerta que “instâncias de URIs podem ter as características de nome (relativas à URN), localizador (relativo à URL) ou de ambos, o que em determinados casos convém a nomeação mais abrangente (URI) do que as específicas”.

Superando esta primeira camada seguiremos para um nível acima, a Camada Sintática.

4.2.2 Camada Sintática (XML e Namespace)

Nesta segunda etapa, nota-se a necessidade de uma estrutura tecnológica que possibilite a representação sintática dos dados de forma a proporcionar a descrição de elementos de maneira clara e simples, podendo ser legível tanto por humanos quanto por máquinas, com a meta de facilitar a elaboração de um complexo processo automatizado de inferências produzidas em conjunto com as tecnologias contidas nas camadas de níveis superiores.

Para tanto, no nível em questão, está inserida a Linguagem eXtensível de Marcação genérica ou *EXtensible Markup Language* – XML, uma linguagem derivada do *Standard Generalized Markup Language* - SGML, tal como HTML (*HyperText Markup Language*, ou em português *Linguagem de Marcação de Hipertexto*), esta última que colaborou em grande escala para o sucesso da *Web*.

A linguagem de marcação extensível propõe descrever as informações tornando-as mais compreensíveis do que o que era feito com HTML, pois ela não possui *tags* pré-determinadas tornando as descrições mais adaptáveis às peculiaridades de cada documento a ser representado.

Um documento XML possui um conjunto de metadados e dados, diferentemente do HTML, ele não se preocupa com a apresentação dos dados e sim

com a definição do seu significado, para que possa ser interpretado independentemente da plataforma e, ainda, servindo de base para linguagens mais complexas da WS, usadas nos níveis seguintes a estes. Segue o exemplo de código na linguagem HTML, figura 9 e do código em XML, figura 10:

Figura 9 – Código em Linguagem HTML

```
<html>

<head>

  <title> Livros sobre Web Semântica </title>

</head>

<body>

  <font face="Arial"><b> A história da Web Semântica</b><br>

    <b><font color="#0000FF">autor 1</font></b>

  <font size="2">2009<br>

</body>

</html>
```

Fonte: Elaborado pela autora.

Figura 10 – Exemplo de código XML.

```
<?xml version="1.0"?>

<acervo>

  <assunto>Livros sobre Web Semântica</assunto>

  <livro ID="1">

    <autor>Autor 1</autor>

    <titulo>A história da web semantica</titulo>

    <ano>2009</ano>

  </livro>

</acervo>
```

Fonte: Elaborado pela autora.

Observando estes códigos podemos notar as semelhanças entre as duas linguagens quanto à sintaxe (a existência dos símbolos “ \leq ” e “ \geq ” entre os metadados e o uso da “/” (barra) para encerrar um bloco); entretanto, a XML, por não conter tags pré-determinadas como o HTML, proporciona grande mobilidade para o desenvolvedor definir descritores que melhor se adequem aos dados por representar, tornando mais clara a manipulação dos dados.

Não se pode deixar a critério de cada profissional a construção das *tags* sem nenhuma forma de controle, pois implicaria também numa forma de incompatibilidade quando os sistemas fossem trocar informações, ideia contrária ao que se prega com o advento da *Web Semântica*.

Nesse sentido, criam-se os *NameSpaces* para discriminar o uso de possíveis terminologias semelhantes para significados diferentes nesta linguagem padrão da W3C.

Os *NameSpaces* são a solução do problema. Um *NameSpace* é uma superetiqueta formada pelo nome da anotação ao qual é concatenado um prefixo. Esse prefixo é definido pelo utilizador e deverá ser único. Para garantir tal unicidade, convencionou-se que se usaria a sintaxe dos URL para o prefixo. Assim, partindo do princípio de que cada utilizador tem um URL próprio, acabam-se os conflitos. (RAMALHO; HENRIQUES, 2001, p. 19-20)

Baseado em Thompson (2004), além da XML há também a *XML SCHEMA*, uma linguagem que define a forma de descrição, (uma “gramática” ou “esquema”) para uma classe de documentos XML, fornecendo elementos para descrever a estrutura e restringir o conteúdo a ser utilizado nos arquivos nesta linguagem.

Podemos compreender a importância da utilização da linguagem XML como padrão para a *Web Semântica*, em função da sua facilidade e clareza ao descrever dados, com o auxílio da linguagem *SCHEMA* e dos espaços de nome. Ela começa a tratar as informações para que se tornem interoperáveis, tornando-se base também para as linguagens mais robustas utilizadas nos níveis superiores.

Outro ponto que torna o XML indispensável é a liberdade oferecida durante a confecção dos documentos, permitindo a cada pessoa construir suas próprias tags. Com este recurso é possível definir o significado dos conteúdos presentes em um documento, processo fundamental para o desenvolvimento de uma estrutura que possua informações processáveis pelos computadores. (CASTELLO BRANCO NETO, 2006, p. 40).

Tal afirmação ratifica a relevância deste nível da arquitetura para tornar objetiva a representação de dados, a fim de prepará-los para adição da semântica nas camadas superiores, para isto avançaremos um nível na estrutura semântica para observar a estrutura adotada na camada RDF.

4.2.3 Camada de dados (*RDF e RDF Schema*)

Após a constatação de que o uso de metadados na *Web* é uma boa alternativa para o processamento inteligente de dados com a estrutura extensível da linguagem XML, nesta camada imediatamente superior à XML adota-se o RDF (*Resource Description Framework*), uma solução baseada em metadados que aplica a linguagem XML e o *Uniform Resource Identifier* – URI para identificar recursos *Web* e propriedades para descrever as relações entre eles.

Diferente da XML que traz metadados com informações sobre partes do recurso informacional, o RDF representa o documento como um todo. Cada recurso representado possui uma estrutura chamada de triplas, onde o recurso é chamado de sujeito, a propriedade é chamada de predicado e o valor de objeto.

Breitman (2005, p. 21-22) nos apresenta um exemplo de código em RDF, conforme figura 11.

Figura 11 – Exemplo de código em RDF.

1. `<?xml version = "1.0"?>`
2. `<rdf: RDF`
3. `xmlns: rdf= "http://www.w3.org/1999/02/22 - rdf-syntax-ns#"`
4. `xmlns: dc= "http://purl.org/dc/elements/1.1/">`
5. `<rdf: Description`
6. `rdf: about = "http://www.inf.puc-rio.br/~karin/index.html">`
7. `<dc: creator>Karin Breitman</dc: creator>`

8. <dc: title>Home Page da Profa. Karin </dc: title>

9. <dc: date> 4 de Outubro, 2004</dc:date>

10. </rdf: Description>

11. </rdf:RDF>

Neste exemplo podemos identificar a base XML utilizada pelo RDF, na linha 1; NameSpaces nas linhas 2 e 3, dentre os elementos característicos do RDF que veremos a seguir:

- Na linha 6, entre aspas, temos um exemplo de **recurso**, este que pode ser qualquer informação que possua uma URI.
- Na linha 7, o elemento em destaque <creator> é uma **propriedade**. Segundo Breitman (2005, p. 22) “uma propriedade é um recurso ao qual possamos nomear, ou seja, o importante será o nome e não o endereço (“o que é” e não “onde está”), embora uma propriedade necessite de um endereço”.
- E por fim, na linha 7 ainda, verificamos que a cadeia de caracteres em vermelho representa o **valor**, este, é o único elemento em RDF que pode conter um recurso, um literal²², valor numérico ou cadeia de caracteres.

Assim, está formado o esquema de triplas (recurso + propriedade + valor) equivalente ao chamado (sujeito + predicado + objeto) para definição de recursos informacionais na *Web*.

A padronização do RDF pelo W3C é repartida em duas informações o próprio RDF – *Resource Description Framework*, como já falamos e o RDF SCHEMA – RDFS, que seguindo a função das linguagens *SCHEMA*, serve para definir padrões para normatizar o uso, neste caso, do RDF, e ainda, possui um vocabulário básico específico para isto, além de mecanismos que propõe a extensibilidade do modelo.

²² Um literal é um valor que é expresso como si mesmo em vez de como um valor variável ou o resultado de uma expressão, como o número 3 ou a sequência "Alô".

Para finalizarmos esta breve observação da camada de dados, veremos que Kleyne e Carrol (2004) destacam alguns motivos que levaram o *RDF* a se tornar uma recomendação para a *Web Semântica*, segundo o *W3C*:

- A capacidade que ele tem de fornecer informações sobre recursos *Web* e os sistemas que o utilizam (Metadados), tais como avaliações de conteúdo, descrições de capacidades e preferências de privacidade;
- A possibilidade de atender aplicações que necessitam de modelos de informações abertos em vez de restritos, para atividades como agendamento, descrição de processos organizacionais e anotações de recursos *Web*;
- Que o *RDF* é capaz de fazer pelas informações processáveis por máquina, o que a *Web* fez pelo hipertexto: permitir que dados sejam processados fora do ambiente particular em que foram criados;
- O seu poder para combinar informações de diferentes aplicações para chegar a novas informações, permitindo o processamento do conteúdo *Web* por agentes de *software* e fornecendo uma linguagem única para eles.

Assim chegamos ao fim desta etapa, onde entendemos um pouco mais sobre a *Web Semântica*, em especial, as três camadas que são o pilar para a elaboração de aplicações nesta estrutura, envolvendo a codificação de caracteres pelo *UNICODE*, agindo em parceria com o identificador universal de recursos na internet, para serem descritos através da linguagem *XML/XML SCHEMA/ NAMESPACES*, que proporcionam a criação de *tags* ou metadados que melhor identifica o documento, servindo de alicerce para a descrição do documento como um todo, através das triplas no *RDF/ RDF SCHEMA*, que age também na camada de dados envolvendo linguagens mais robustas, como o *SPARQL*, mostradas na figura 5.

Veremos que as bibliotecas digitais também estão fazendo uso de tecnologias para melhor descrição de dados e auxílio na interoperabilidade de recursos informacionais na *Web*, afinal a questão do intercâmbio e disseminação de informações são questões de cunho fundamental para a Biblioteconomia, sendo que muitas destas tecnologias utilizadas pertencem às três camadas supracitadas, o que pode aproximar cada vez mais as bibliotecas digitais da arquitetura da *Web Semântica*.

4.3 O que as Bibliotecas Digitais têm em comum com as camadas-base da Web Semântica?

Vimos anteriormente que os três primeiros níveis da arquitetura da *Web Semântica* dirigem seus esforços para a representação descritiva dos dados, para que eles possam ser acessíveis pelo maior número de pessoas e máquinas tentando dirimir barreiras de incompatibilidade entre plataformas, língua etc.

Como um sistema fechado, as bibliotecas digitais também podem refletir sobre a implantação destes quesitos, por possuírem um modelo de estrutura colaborativa e altamente padronizado, auxiliando na organização da *Web Semântica*.

Na implantação de bibliotecas digitais deve-se ter atenção a uma série de situações para que ela funcione da melhor maneira e de modo que atenda as expectativas do usuário, quesitos como conhecimento sobre direitos autorais, infraestrutura tecnológica, representação e interoperabilidade de dados, visando com que os mesmos não se prendam a qualquer plataforma para que não se tornem um empecilho para visualização em diferentes sistemas operacionais e *browsers*.

As recomendações feitas para *Web*, pelo *World Wide Web Consortium*, que tem por objetivo trazer maior organização, padronização e semântica para as informações na rede, em muito pode ter sido inspiradas nos métodos advindos da Biblioteconomia e ao contrário, as práticas da *Web Semântica* podem ajudar nos padrões tecnológicos que objetivam as transações de dados entre diferentes sistemas.

Vejamos então alguns padrões utilizados pelas bibliotecas para a representação de metadados, na estruturação e descrição dos dados para a modelagem do ambiente informacional.

4.3.1 O Formato MARC 21

O padrão MARC surgiu da necessidade da participação de bibliotecas em redes e em programas cooperativos de catalogação que requer a adoção de padrões e a aplicação de normas, produzindo registros descritivos confiáveis e de qualidade; que possibilitem o intercâmbio de dados, no âmbito universal; e a interoperabilidade entre os sistemas de informação. (PEREZ, 2008).

A partir da fusão do USMARC (*Machine Readable Cataloging*) dos Estados Unidos, criado pela Biblioteca do Congresso Americano (LC), com o objetivo de adotar um padrão internacional para a descrição bibliográfica, com o MARC canadense, surgiu o MARC 21, o formato para intercâmbio de dados, ou segundo sua sigla, cujo significado é “catalogação legível por máquinas” amplamente manipulado na atualidade.

É importante ressaltar três aspectos da estrutura de descritores do formato MARC 21: as etiquetas, os indicadores e os subcampos, destacados na figura 12:

Figura 12 – Exemplo de MARC 21

Líder 01041cam 2200265 a 4500
Nº. de Controle 001 ###89048230
Iden. Nº. Controle 003 DLC
Ult. Intervenção 005 19911106082810.9
Campos Fixos 008 891101s1990 maua j 001 0 eng
Nº. Controle LC 010 ## \$a ###89048230
ISBN 020 ## \$a 0316107514 : \$c \$12.95
ISBN 020 ## \$a 0316107506 (pbk.) : \$c \$5.95 (\$6.95 Can.)
Fonte de Catalog. 040 ## \$a DLC; \$c DLC; \$d DLC
Nº Chamada LC. 050 00 \$a GV943.25; \$b .B74 1990
Classif. Dewey 082 00 \$a 796.334/2; \$2 20
Nome Pessoal 100 **1#** \$a Brenner, Richard J., ; \$d 1941-
Título 245 10 \$a Make the team. \$p Soccer : \$b a heads up guide to
 super
 soccer! / \$c Richard J. Brenner.
Título Variante 246 30 \$a Heads up guide to super soccer
Edição 250 ## \$a 1st ed.
Imprensa 260 ## \$a Boston : \$b Little, Brown, \$c c1990.
Desc. Física 300 ## \$a 127 p. : \$b ill. ; \$c 19 cm.
Nota Geral 500 ## \$a "A Sports illustrated for kids book."

Nota Resumo 520 ## \$a Instructions for improving soccer skills.
 Discusses dribbling,
 heading, playmaking, defense, conditioning, mental attitude, how to
 handle problems with
 coaches, parents, and other players, and the history of soccer.
Assunto: 650 #0 \$a Soccer \$v Juvenile literature.
Assunto: 650 #1 \$a Soccer.

Fonte: (Siqueira, 2003, p. 45)

Neste trecho de código conforma a figura 12, visualizamos o que usualmente chama-se de etiquetas no MARC, que são os três números que iniciam a descrição da informação (001, 008, 100, 245 [...]) e cada um tem uma atribuição específica, para exemplificar, a etiqueta 245 representa o título da obra.

Após as etiquetas encontramos os indicadores simbolizados por dois caracteres (números, caracteres especiais ou ambos), como o indicador em destaque na linha 12, posterior à etiqueta 100, #1. Ressalta-se que há campos (etiquetas) como 008, onde existe uma sequência numérica que não é considerada “indicador”.

Por fim, depois dos indicadores existe um símbolo (\$) chamado de subcampo seguido da letra (a), destacados na última linha, no campo 650, para melhor visualização.

Com isto, vimos de modo bastante sucinto como é feita a representação dos dados bibliográficos por meio do formato MARC 21, bastante difundido no ambiente biblioteconômico, entretanto, com a explosão informacional digital necessita-se de padrões com maior flexibilidade com uma estrutura que possa dar cobertura a estes documentos. Flamino e Santos (2006, p.3) confirmam,

[...] que o formato de intercâmbio MARC (Machine Readable Cataloging) tem permitido às instituições o intercâmbio de dados bibliográficos e catalográficos por décadas, favorecendo o acesso aos conteúdos informacionais contidos em diversos acervos. Mas, o crescimento exponencial de informações e da geração de documentos, sobretudo em ambiente digital, tem exigido uma maior flexibilidade e interoperabilidade entre os sistemas informacionais disponíveis.

Para tanto, alguns autores pregam o breve desuso do formato MARC 21, tais como Tennant (2004) em seu texto intitulado “*MARC must die*”. O que se pode notar é a modernização do padrão que tenta se atualizar integrando-se à linguagem XML, ao invés de ser substituído por ela. Yee (2004, p. 7) segue em defesa do MARC, contradizendo autores de outra linha, quando rebate críticas sobre a estrutura e redundância existentes no formato. Elucida que a maioria dos problemas associados ao MARC são na verdade, dificuldades com a catalogação e as “redundâncias” seguem a propósitos distintos.

Na tentativa de moldar-se às realidades mais exigentes em questões de interoperabilidade e flexibilidade, o MARC 21 agrega elementos da linguagem XML formando o padrão MARCXML, o que ratifica a defesa de Yee (2004) e podemos assim encerrar esta breve “conversa” sobre este padrão de metadados com a colocação de Moreno e Brascher (2007, p.19), “no entanto, o uso do formato ainda é a saída viável para aproveitamento dos dados já registrados e a transição para o MARCXML é um reconhecimento da modernidade ou atualidade deste”.

Dessa forma, verificamos que os padrões da Biblioteconomia e da Ciência da Informação podem contribuir com o desenvolvimento da *Web Semântica*, por oferecerem estruturas para a descrição dos dados altamente padronizadas.

4.3.2 O padrão de metadados Dublin Core

Outro padrão de metadados que merece ser citado é o Dublin Core, um padrão simples composto por um grupo básico de 15 elementos descritores, fazendo disto um ponto forte, pois é claro e objetivo, mas também frágil, por não acomodar uma semântica mais expressiva. No quadro 4 estão os elementos básicos mencionados:

Quadro 4 – Elementos básicos do padrão Dublin Core

Assunto (<i>subject</i>)	Tópico abordado pelo trabalho
Título (<i>title</i>)	Nome do objeto
Criador (<i>creator</i>)	Responsável(eis)pelo conteúdo

	intelectual do objeto
Descrição (<i>description</i>)	Descrição do conteúdo do objeto
Editor (<i>publisher</i>)	Agente ou agência responsável por disponibilizar o objeto
Outro agente (<i>contributor</i>)	Pessoa(s) que fez (fizeram) contribuições significativas para o objeto
Data (<i>date</i>)	Data da publicação
Tipo de objeto (<i>type</i>)	Gênero do objeto, se ficção, novela, poema ou dicionário
Formato (<i>format</i>)	Manifestação física do objeto. Ex.: arquivos executáveis, do tipo texto ou PDF
Identificador (<i>identifier</i>)	Cadeia ou número utilizado para identificar unicamente aquele objeto
Relacionamento (<i>relation</i>)	Relacionamento com outros objetos
Fonte (<i>source</i>)	Outros objetos, eletrônicos ou físicos, dos quais este foi derivado (caso seja aplicável)
Linguagem (<i>language</i>)	Linguagem do conteúdo intelectual
Cobertura (<i>coverage</i>)	Localizações espaciais e durações temporais características do objeto
Direito (<i>rights</i>)	Informação sobre os direitos acerca do objeto

Fonte: Breitman (2005, p. 18).

O *Dublin Core Metadata Initiative*- DCMI²³ é uma iniciativa dos usuários do padrão e a partir da perspectiva desta comunidade os metadados estão caracterizados em quatro níveis de interoperabilidade, segundo informado em seu site:

- **Nível 1 (compartilhados definições de termos).** A interoperabilidade entre metadados usando aplicativos é baseada na partilha de definições da linguagem natural. Dentro de um ambiente de aplicação, como uma intranet, sistema de biblioteca, ou repositório federação, os participantes concordam com os termos para usar em seus metadados e como esses termos são definidos. Termos são adotados em aplicações utilizando tecnologias específicas de implementação e interoperabilidade com o “resto do mundo”.

- **Nível 2 (interoperabilidade semântica formal).** A interoperabilidade entre metadados usando aplicativos é baseada no modelo compartilhado formal proporcionado por RDF, que é utilizado para apoiar *Linked Data*. Conforme definido na *Wikipedia*, o termo "Linked Data" descreve "uma prática recomendada para expor, compartilhar peças de ligação de dados, informações e conhecimento sobre a *Web Semântica* utilizando [endereços Web] URIs e RDF." As propriedades e classes de termos de metadados DCMI foram definidos para a compatibilidade com os princípios *Linked Data*. Nos últimos dois anos, grandes quantidades de dados comerciais e do setor público foram adicionados a uma nuvem de dados de crescimento vinculado. Os motores de busca como Yahoo e gerenciamento de conteúdo, plataformas como Drupal implementaram suporte para RDFa, um método para expor dados ligados embutidos em páginas da *Web*. Com efeito, a ideia de fundação de Dublin Core - "metadados simples para descoberta de recursos" - está sendo reinventada sob a bandeira de "dados estruturados para a otimização de motor de busca". Dos quatro níveis de interoperabilidade, esta parece estar crescendo mais rapidamente.

- **Nível 3 (Descrição interoperabilidade Set sintática).** Os aplicativos são compatíveis com o modelo de dados ligados e, além disso, partes de uma

²³ <http://dublincore.org>

sintaxe abstrata para registros de metadados comprováveis, o "conjunto de descrição".

- **Nível 4 (Descrição interoperabilidade Perfil Set).** Os registros trocados entre metadados usando aplicativos com um conjunto de restrições, adotando-se os mesmos vocabulários, e refletem um modelo compartilhado do mundo.

Os níveis 3 e 4 são mais experimentais do que os níveis 1 e 2, na medida que eles não são tão bem apoiados com instrumentos de *software*, embora os problemas abordados neste trabalho devem crescer em importância, como produtores de registros de metadados que devem mover suas informações em um ambiente ligado de dados .

Os padrões apresentados aqui foram apenas alguns dentre tantos existentes, que viabilizam o intercâmbio de informações, com diferentes finalidades específicas e estão mais relacionados às bibliotecas, repositórios digitais etc. O DC, por exemplo, “é um padrão voltado principalmente para a descrição e a descoberta de documentos eletrônicos”. (MARCONDES, 2008, p.108).

A literatura científica aponta que há críticas relativas à robustez de padrões como o MARC 21, devido à falta de flexibilidade e o Dublin Core em virtude da simplicidade excessiva.

O importante é que independentemente do uso ou desuso de determinados padrões em bibliotecas digitais, linguagens de descrição como o XML parecem se destacar em meio a este “mar de ferramentas” que trabalham em função da manipulação de dados e metadados, no intuito de organizar o ambiente digital, assim como o RDF e a utilização de URIs.

Com isto é possível à percepção de que o sistema de organização utilizado nas bibliotecas mesmo antes da sua presença no meio digital pode ter inspirado as formas de arranjo estudadas atualmente para as aplicações consideradas na *Web Semântica*.

XML é, cada vez mais, a base de vários padrões da área de informação. Além do OAI-PMH, outro protocolo de recuperação de informações, o SRW – “Search Retrieval on Web”, o sucessor do protocolo Z39.50, de largo uso em bibliotecas para viabilizar a consulta a um catálogo virtual formado por catálogos distribuídos de várias bibliotecas, tem por base para troca de mensagens entre programas clientes e servidores XML. As ontologias, bases de

conhecimentos sobre determinado domínio, baseiam-se na OWL – “Ontology Web Language”, codificada em XML. [...] A base da infraestrutura da Web Semântica vão ser páginas escritas em XML, que, além do conteúdo, terão metadados utilizando vocabulários e relações muito poderosas (utilizando RDF e ontologias) para expressar a semântica das novas páginas *Web*. (MARCONDES, 2008, p.108-109).

Em contrapartida, as tecnologias indicadas para a composição dos níveis básicos da arquitetura desta *Web*, se encontram também recomendadas ou já utilizadas em bibliotecas digitais, o que proporciona um alicerce notável para inserção das bibliotecas no ambiente digital semântico.

Uma biblioteca semântica, que pode atribuir significado aos descritores de metadados e dados, traz uma série de benefícios, os quais se destacam a recuperação de informações mais precisas e significativas nas solicitações de busca pelo usuário final.

Para tanto, não é preciso abandonar ou deixar de adquirir os elementos associados à *Web* 2.0, pois como foi explicado no capítulo 2, a *Web* está em constante processo evolutivo, onde em determinados períodos de tempo predomina um conjunto de práticas tecnológicas, mas que não excluem totalmente as práticas anteriores. Um exemplo disto é a Biblioteca Digital Semântica Jerome DL, que diz mesclar itens da *Web* Social com a *Web* Semântica.

A JeromeDL é uma Biblioteca Digital Semântica Social. Como uma biblioteca digital que permite às instituições publicar facilmente documentos *Web*, ela suporta uma variedade de formatos e permite armazenar e consultar uma rica descrição bibliográfica de cada documento. Para encontrar documentos relevantes, usuários da JeromeDL podem usar recursos de pesquisa e navegação, além de poder permitir que outros vejam seus marcadores e anotações e também compartilhar seus conhecimentos dentro de uma rede social. A JeromeDL pode ainda tratar um recurso informacional como um *post* de blog, permitindo aos usuários comentar o conteúdo do recurso e responder aos comentários dos outros e, dessa forma, criar novos conhecimentos. (DIGITAL ENTERPRISE RESEARCH INSTITUTE; GDANSK UNIVERSITY OF TECHNOLOGY, 2009).

Nesta pesquisa procurou-se conhecer através da linha evolutiva da *Web* como se chegou à tão comentada *Web* Semântica, com suas questionadas nomenclaturas, e assim contextualizá-la com as bibliotecas digitais.

Ampliamos nossos esforços para o estudo das tecnologias empregadas nas três camadas mais básicas da arquitetura da *Web Semântica*, para associar aos padrões que são utilizados nas bibliotecas digitais tentando traçar um paralelo entre eles, cujo foco principal é a melhor maneira para representar e descrever dados e metadados entre distintos sistemas sem perdas de informação via rede de computadores.

Pretendemos com isto incentivar a adoção de métodos de desenvolvimento semântico para as bibliotecas digitais, tendo em vista, que a base da infraestrutura já está elaborada, conforme as semelhanças de finalidades e o uso de padrões que se adaptam aos recursos tecnológicos das camadas e que, num breve futuro, possamos ter mais bibliotecas semânticas trazendo o que há de potencial nas tecnologias *Web* com aplicações no ambiente das bibliotecas digitais.

5 CONSIDERAÇÕES FINAIS

Foi possível através do estudo realizado conhecer um pouco da linha evolutiva da *Web* e da inserção das bibliotecas neste âmbito, através de demonstrações das tecnologias nas camadas iniciais da estrutura semântica e alguns padrões de metadados utilizados nas bibliotecas digitais. Recomenda-se que a representação por metadados em bibliotecas digitais seja perfeitamente aplicável à *Web Semântica*, e o potencial das tecnologias computacionais aplicados nas bibliotecas digitais.

A integração de tecnologias como a linguagem XML ao MARC 21 mostra o esforço da comunidade biblioteconômica que o desenvolveu em adaptar este padrão às formas mais flexíveis de desenvolvimento de ambientes informacionais digitais contribuindo com o estabelecimento da *Web Semântica*. O mesmo ocorre com o Dublin Core, entretanto, com uma representação mais simplificada e que muitas vezes não atende a uma lógica de descrição estabelecida na Biblioteconomia.

As bibliotecas digitais com seus padrões específicos podem fazer uso das técnicas que as tornem mais semânticas; os primeiros passos para a concretização desta ideia estão na adaptação às camadas analisadas da arquitetura da *Web Semântica*, onde já estamos em caminhos sólidos de semelhanças.

O intercâmbio de informações com outros sistemas que não fossem de bibliotecas seria facilitado, pois pelo que se pode perceber pela ideia de Berners-Lee (2005), é conectar todos os sistemas de modo que haja a troca de documentos sem perda de dados.

O primeiro passo para o intercâmbio de dados está na camada Unicode/ URI, uma camada bastante específica que se preocupa com a padronização de caracteres e a localização de recursos informacionais digitais. Esta questão é importante para a interoperabilidade dos dados com o objetivo de que não se perca nenhum caractere no intercâmbio de documentos, além de que tenham um endereço que os identifique como um todo, o que é fundamental para o intercâmbio de recursos numa biblioteca em rede.

O segundo passo está na camada XML, que propõe a linguagem XML como padrão para a descrição de dados, por ser clara e flexível e amplamente utilizada em diversos sistemas. Já nas bibliotecas digitais existe o padrão MARC21 e também a

fusão MARCXML com a mesma proposta de representação descritiva, ratificando as semelhanças na finalidade e agregação de instrumentos computacionais.

O terceiro passo, da camada RDF, que visa identificar o conteúdo do documento, o teor de que se trata, neste sentido há o Dublin Core que apesar de simples contempla a mesma situação e também pode ser agregado ao RDF.

Deste modo podemos perceber que as bibliotecas digitais tem uma base favorável para adaptar-se às exigências da *Web Semântica* e isto não é somente algo teórico, como é o caso da Biblioteca Digital Social Semântica – Jerome- DL, que já existe e agrega características da *Web 2.0*.

Com esta pesquisa tentou-se demonstrar a integração dos instrumentos tecnológicos atrelados à *Web 3.0* com as unidades de informação, em ambiente virtual, para que deste modo, possamos enxergar neste início, uma porta para o desenvolvimento de bibliotecas digitais semânticas, além da associação do profissional bibliotecário às Tecnologias da Informação e Comunicação, com intuito de estimular pesquisas mais específicas nesta linha.

Fica comprovada a troca de informações entre tecnologias e práticas da Ciência da Informação e da Computação, reforçando que as bibliotecas digitais, pautadas nas práticas biblioteconômicas fornecem subsídios para a elaboração de novas tecnologias no domínio *Web*, e por sua vez, as tecnologias computacionais trazem novos recursos para a efetivação destas práticas em bibliotecas digitais, potencializando a busca e a recuperação aos usuários dos sistemas e ambientes informacionais.

Assim, é possível identificar que há uma sinergia entre bibliotecas digitais e *Web Semântica*, pois a primeira contribui com seu ambiente padronizado, pautado nas metodologias da Biblioteconomia, tais como a representação descritiva; a segunda oferece nas tecnologias computacionais semânticas, a possibilidade de potencializar os serviços oferecidos pelas bibliotecas digitais, potencializando a busca e agregando maior semântica aos recursos informacionais, facilitando a recuperação pelos agentes computacionais e apresentando informações consistentes e significativas para o usuário final.

Como pesquisas futuras sugere-se o estudo da arquitetura da *Web Semântica* na sua integridade, além de aprofundamentos em alguns protocolos de recuperação de informação, padrões de interoperabilidade e diversas outras tecnologias inerentes também às bibliotecas digitais (RDA, OAI-PMH, SRU/SRW, Z39.50 etc.),

considerados elementos fundamentais na constituição e na modelagem de ambientes informacionais digitais contemporaneamente, pois entende-se que desta forma os bibliotecários possam participar cada vez mais das discussões sobre as tecnologias da informação emergentes no ambiente *Web*.

Ainda nesta linha ratifica-se a necessidade do bacharel em biblioteconomia, na condição de profissional da informação estar engajado com a pesquisa científica para que deste modo possa trazer os questionamentos sobre assuntos relacionados à informação, sobretudo na *Web*, pondo esta questão em pauta também no meio biblioteconômico de forma mais expressiva do que há atualmente.

Conforme o que foi expresso durante esta pesquisa existem semelhanças entre a prática da biblioteconomia e a informática no que cerne ao tratamento da informação em rede, principalmente, entretanto a participação da biblioteconomia na descoberta de métodos que visem a tratar os dados e metadados com a finalidade de recuperar recursos informacionais com robustez, parece ser algo ainda dissociado do fazer bibliotecário pela sociedade em geral e as vezes até pelo próprio profissional, o que acaba por restringir o ambiente de atuação dos egressos da área.

Neste sentido conhecer as próprias competências e ampliar interdisciplinarmente os horizontes, faz com que o profissional possa enxergar a contribuição da biblioteconomia para inúmeras áreas que necessitam do tratamento, busca e recuperação da informação. Ganhada esta segurança, o passo seguinte para que haja maior participação na sociedade é a divulgação desta descoberta, para contribuir efetivamente com as outras áreas e permitir a inclusão destas na biblioteconomia, compondo uma colaboração mútua como acontece entre as Ciência da Informação e da Computação.

Por fim uma postura profissional ativa trará aos bibliotecários maior reconhecimento de um trabalho extremamente relevante que é exercido há décadas e está sempre se renovando junto com as Tecnologias da Informação e Comunicação, o que se leva a concluir que o bibliotecário nunca se torna um profissional arcano e sim atual e dinâmico que ainda tem muito a colaborar com a massa documental crescente nos diferentes suportes informacionais, em especial na *Web*, que é o foco deste trabalho.

REFERÊNCIAS

- ABRAM, S. 15 Minutes a day: a personal learning management strategy. **Info Tech Column**, 20 Dec. 2006. Disponível em: <https://docs.google.com/viewer?a=v&q=cache:QllpZliE2AYJ:stephenslighthouse.com/files/IOColumn_56.pdf+&hl=ptBR&gl=br&pid=bl&srcid=ADGEESgg9ttH7PocViy2QIZ7a5OgvEfique7nRB1akYBBqKw0rfzds0Qt7GLkg72DfgrOZc9UNzHzJA4g_hYFpAlkELkHjFqEbaevmewkhwUDcm6K27KL97iw4v0FEnmGGEQ50sesr2n&sig=AHIEtbTaUu0-HrGuqLEjFPPs_IDHv9Y_wA>. Acesso em: 03 jan. 2013.
- AGHAEI, S.; NEMATBAKHSH, M. A.; FARSANI, H. K. Evolution of the World Wide Web: Web 1.0 to Web 4.0. **International Journal of Web & Semantic Technology (IJWesT)** v. 3, n. 1, Jan. /2012. Disponível em : < <http://www.airccse.org/journal/ijwest/papers/3112ijwest01.pdf> >. Acesso em: 16 fev. 2013.
- ALVARENGA, L. A teoria do conceito revisitada em conexão com ontologias e metadados no contexto das bibliotecas tradicionais e digitais. **DataGramZero**, v. 2, n. 6, dez./2001. Disponível em: < http://dgz.org.br/dez01/Art_05.htm >. Acesso em: 22 fev. 2013.
- ALVES, R.C. V. **Web Semântica**: uma análise focada no uso de metadados. 2005.180 f. Dissertação (Mestrado em Ciência da Informação)) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. **The Semantic Web**: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, May. 2001.
- _____.; et al. Uniform Resource Identifier (URI): Generic Syntax, jan. 2005. The Internet Society. Disponível em: < [ttp://tools.ietf.org/pdf/rfc3986.pdf](http://tools.ietf.org/pdf/rfc3986.pdf)>. Acesso em: 28 fev. 2013.
- BLATTMANN, U.; SILVA, F. C.C. da. Colaboração e interação na web 2.0 e biblioteca 2.0. **Revista ACB: Biblioteconomia em Santa Catarina**, v. 12, n. 2, p. 191-215, 2007.
- BORGES, Maria Manuel. Biblioteca digital: materialização e utopia.**Revista da Faculdade de Letras: Ciências e Técnicas do Patrimônio**, v. 2, p. 653-664, 2003.
- BREITMAN, K. **Web Semântica**: a internet do futuro. Rio de Janeiro: LTC, 2006. 190 p.

BULCÃO NETO, R. de F.; PRAZERES, V. S.; PIMENTEL, M. da G. C. **Web Semântica: conceitos e práticas** – 2006. Disponível em <<http://homes.dcc.ufba.br/~prazerres/papers/BulcaoEtAlli-WebMedia2006Course.pdf>>. Acesso em 22 fev. 2013.

BUSH, V. As We May Think. **The Atlantic** . 1945. Disponível em: <<http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/4/>>. Acesso em: 10 mar. 2013.

CASTRO, F. F. de. **Elementos de interoperabilidade na catalogação descritiva: configurações contemporâneas para a modelagem de ambientes informacionais digitais**. 2012. 202 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.

_____. **Padrões de descrição e representação de recursos informacionais em bibliotecas digitais na perspectiva da Ciência da Informação: uma abordagem do MarcOnt Initiative na era da Web Semântica**. 2008.196 f. Dissertação (Mestrado em Ciência da Informação)) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.

CASTELLO BRANCO NETO, W. **Web Semântica na construção de sistemas de aprendizagem adaptativos**. 2006. 215f. Tese (Doutorado em Ciência da Computação) – Universidade Federal de Santa Catarina, Santa Catarina, 2006.

CHOI, Y.; RASMUSSEN, E. What is needed to educate future digital librarians. **D-lib magazine**, v. 12, n. 9, p. 1-8. Sept. 2006. Disponível em: <<http://www.dlib.org/dlib/september06/choi/09choi.html>>. Acesso em: 28 fev. 2013.

CORRÊA, C. H. W. Arquitetura participativa da internet: social software e web 2.0. In: Gustavo Cimadevilla. (Org.). **Comunicación, Tecnología y Desarrollo. Trayectorias**. Río Cuarto: Universidad Nacional de Río Cuarto, 2008, v. 1, p. 187-202.

COUTINHO, C. P.; BOTTENTUIT JUNIOR, J. B. Comunicação educacional: do modelo unidireccional para a comunicação multidireccional na sociedade do conhecimento. Comunicação e Cidadania. In: **Actas do 5º Congresso da Associação Portuguesa de Ciências da Comunicação**. Braga: Centro de Estudos de comunicação e sociedade Universidade do Minho, 6-8 set. 2008.

CORMODE, G.; KRISHNAMURTHY, B. Key differences between Web 1.0 and Web 2.0. **First Monday**. v. 13. n. 6. 2 June 2008. Disponível em: <<http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2125/1972>>. Acesso em: 22 jan. 2013.

COYLE, K. The library catalog in a 2.0 world. **Journal of Academic Librarianship**, v. 33, n. 2. 2007. Disponível em < http://www.kcoyle.net/jal_33_2.html>. Acesso em: 22 jan. 2013.

CUNHA, M. B. da. Das bibliotecas convencionais às digitais: diferenças e convergências. **Perspectivas em Ciência da Informação**, v. 13, n. 1, p. 2-17, jan./abr. 2008. Disponível em: < <http://www.scielo.br/pdf/pci/v13n1/v13n1a02.pdf>>. Acesso em: 26 fev. 2013.

_____. Desafios na construção de uma biblioteca digital. **Ciência da Informação**, v. 28, n. 3, p. 257-268, set./dez. 1999. Disponível em: < <http://www.scielo.br/pdf/ci/v28n3/v28n3a3.pdf> >. Acesso em: 11 mar. 2013.

_____.; MCCARTHY, C. Estado atual das bibliotecas digitais no Brasil. In: Carlos H. Marcondes; Helio Kuramoto; Lidia Brandão Toutain; Luis Sayão. (Org.). **Bibliotecas digitais: saberes e práticas**. 2. ed. Brasília: IBICT, 2008, p. 25-54.

DAVIS, I. "Talis, Web 2.0 and All That", **Internet Alchemy** blog, 4 July, 2005. Disponível em < <http://internetalchemy.org/2005/07/talis-web-20-and-all-that>>. Acesso em: 25 jan. 2013.

DIAS, T. D; SANTOS, N. Web Semântica: conceitos básicos e tecnologias associadas. **Cadernos do IME**, Rio de Janeiro, v. 14, n. 7. jun 2003. Disponível em <<http://www.ime.uerj.br/cadernos/cadinf/vol14/7-neide.pdf>>. Acesso em 05 mar 2013.

DIGITAL ENTERPRISE RESEARCH INSTITUTE; GDANSK UNIVERSITY OF TECHNOLOGY. **Jerome DL – e-Library with semantics**. 2004-2009. Disponível em:<http://www.jeromedl.org/index.php?option=com_content&task=view&id=39&Itemid=94>. Acesso em: 11 mar. 2013.

FLAMINO, A. N.; SANTOS, P. L. V. A, da C. MARCXML para OAI. In: **XIV SEMINÁRIO NACIONAL DE BIBLIOTECAS DIGITAIS (SNBU)**, 2006, Salvador. Disponível em:<https://docs.google.com/viewer?a=v&q=cache:rwgM8Vdktm4J:eprints.rclis.org/16622/1/FLAMINO_AN_MARCXML_para_a_OAI.pdf+&hl=pt-BR&gl=br&pid=bl&srcid=ADGEESgxOCOBVqw1MzotU0f2zUop3i89rpTY84ScRjcL4mIYVOuWd7ldskykuYw2BotT4JS0pYDCe5kEJGZj38_Mvv23TuG3RpqR7GjvrLaAUXpm55FfZTFiq_uSUVGo-c03EDCxcQsw&sig=AHIEtbQPSJIhx_6bS3w27_ByGhp_qvh-aA>. Acesso em: 03 fev. 2013.

FRANCISCATO, F. T. **ROAD: Repositório Semântico de Objetos de Aprendizagem para Dispositivos Móveis**. 2010.101f. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Santa Maria, Rio Grande do Sul. 2010.

FONSECA, E. N. **Introdução à Biblioteconomia**. 2. ed. Brasília: Briquet de Lemos Livros, 2007. 152 p.

GOÑI, Jorge L.; FERNANDES, M. C. P.; LUCENA, Carlos J. P. de. E-Learning e a Web Semântica. **PUC-RIOInf.MCC**, Rio de Janeiro, v.2, n.12, p. 1-18 jun. 2002. Disponível em < ftp://ftp.inf.puc-rio.br/pub/docs/techreports/02_12_goni.pdf>. Acesso em 10 mar 2013.

GUIMARÃES, H. B. A. **A Web 2.0 na Dinamização de um Portal Educativo**. 2009. 79 f. Dissertação (Mestrado em Tecnologia Multimídia). Universidade do Porto, Porto.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. De 2005 para 2008, acesso à internet aumenta 75,3%. IBGE, dez. 2009. Disponível em: <<http://saladeimprensa.ibge.gov.br/noticias?view=noticia&id=1&busca=1&idnoticia=1517>>. Acesso em: 06 jan. 2013.

ISHIDA, R. Introdução a conjuntos de caracteres e codificações. MIT, ERCIM, Keio, Beihang.: **W3C**. 2010. Disponível em: <<http://www.w3.org/International/getting-started/characters>>. Acesso em: 05 fev. 2013.

KLYNE, G.; CARROLL, J. J. Resource Description Framework (RDF) Concepts and Abstract Syntax. W3C, feb. 2004. Disponível em: <<http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/>>. Acesso em: 13 mar. 2013.

LÉVY, P. **O que é virtual?**. São Paulo: OCR BR, 2000. 110 p

_____. **Cibercultura**. São Paulo: Editora 34. 2. ed. 251 p.

_____. **Inteligência coletiva**. São Paulo: Edições Loyola, 1998. 212p

MARCHIORI, P. Z. “Ciberteca” ou biblioteca virtual: uma perspectiva de gerenciamento de recursos de informação. *Ciência da Informação*, v. 26, n.2. 1997. Disponível em: < <http://revista.ibict.br/ciinf/index.php/ciinf/article/view/389/349>>. Acesso em: 13 fev. 2013.

MARCONDES, C. H. Em busca de uma semântica do digital ou “As They May Think”. **Ponto de Acesso**, Salvador, v.6, n.2, p. 35-73, dez. 2012. Disponível em < <http://www.portalseer.ufba.br/index.php/revistaici/article/view/6103/4561> >. Acesso em 13 jan 2013.

_____; KURAMOTO, H.; TOUTAIN, L. B.; SAYÃO, Luis F.(Org.). **Bibliotecas digitais: saberes e práticas**. Salvador: UFBA; Brasília: IBICT, 2008. p. 15-25.

MILANESI, L. **Biblioteca**. São Paulo: Ateliê Editorial, 2002. 115 p.

MORENO, F. P.; BRASCHER, M. MARC, MARCXML e FRBR: relações encontradas na literatura. **Inf. & Soc.:Est.**, João Pessoa, v.17, n.3, p.13-25, set./dez. 2007. Disponível em: <<http://www.pergamum.pucpr.br/redepergamum/trabs/2008.pdf>>. Acesso em 01 mar. 2013.

NOVAES, S. C. **Imagem, magia e imaginação: desafios ao texto antropológico**. *Mana* [online]. 2008, vol.14, n.2, pp. 455-475. ISSN 0104-9313.

OLIVEIRA, M. L. A. de. **Folksonomia: uma indexação livre e social das informações na web**. 2008. 53f.. Monografia (Bacharelado em Biblioteconomia) – Universidade Federal do Rio Grande do Norte, Natal. 2008.

O'REILLY, T. What is Web 2.0 Design Patterns and Business Models for the next generation of software. **Communications & Strategies**. n. 1, p. 17, First Quarter 2007. Disponível em: <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1008839>. Acesso em: 11 jan. 2013.

ORTEGA, C. D. Relações históricas entre Biblioteconomia, Documentação e Ciência da Informação. **Revista de Ciência da Informação**. v. 5. n. 5, out./2004. Disponível em: <http://dgz.org.br/out04/Art_03.htm>. Acesso em: 10 mar. 2013.

PEREIRA, M. de N. F.; PINHEIRO, L. V. R. (Org.). **O sonho de Otlet: aventura em tecnologia da informação e comunicação**. Rio de Janeiro; Brasília: IBICT/DEP/DDI, 2000. 291 p.

PEREIRA, E. C.; RUTINA, R. O século XXI e o sonho da biblioteca universal: quase seis mil anos de evolução na produção, registro e socialização do conhecimento. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 4, n.1, p. 5-19, jan./jun. 1999. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/590/359>>. Acesso em 01 fev. 2013.

PEREZ, D. R. MARC 21: apresentação. Rio de Janeiro, out. 2008. Disponível em: <<http://www.dbd.puc-rio.br/MARC21/conteudo.html>>. Acesso em 12 mar. 2013.

PICKLER, M. E. V. Web Semântica: ontologias como representação do conhecimento. **Perspectivas em Ciência da Informação**. v. 12. n.1. p.65-83, jan./abr. 2007. Disponível em: <<http://www.scielo.br/pdf/pci/v12n1/05.pdf>>. Acesso em: 05 dez. 2012.

PRIMO, A.. O aspecto relacional das interações na Web 2.0. **E- Compós** (Brasília), v. 9, p. 1-21, 2007. Disponível em: <
http://www.moodle.ufba.br/file.php/10203/cultura_digital/web2_primo.pdf >. Acesso em 02 jan. 2013.

RAMALHO, J. C. L.; HENRIQUES, Pedro R. **XML e XSL: da teoria à prática**. Caxias do Sul: FCA. 2001. 100 p.

RAMALHO, R. A. S.; VIDOTTI, S. A. B. G.; FUJITA, Mariângela S. L. Web semântica: uma investigação sob o olhar da Ciência da Informação.

DataGramZero, v. 8, n. 6, dez. 2007. Disponível em <
http://www.datagramazero.org.br/dez07/Art_04.htm>. Acesso em 13 fev. 2013.

RIBEIRO, F. **Biblioteca: novos termos para um velho conceito**. Porto: Universidade do Porto. 1996. 100 p.

RODRIGUES JUNIOR, F. Unicode. 2011. Disponível em: <
https://docs.google.com/viewer?a=v&q=cache:NeJYpYHUJb0J:lcadfs2.lcad.icmc.usp.br/~junio/mat/utf.pdf+&hl=pt-BR&gl=br&pid=bl&srcid=ADGEESjmBGHcdbZxiORA7C-s840RRw_c3T6A91L4WyPt5EQNUq21N3tmY8f5iGAJySmytV221i7g8jGn47rEAYJjR34SzVsgq0yjjNTd6_6Nd8zQq2Q9YEpHfBS9_Eg4FzpTJBL1JZAe&sig=AHIEtbQxGn84RkoFCz8MT67Dycor9RMOWw >. Acesso em: 11 jan. 2013.

ROSETTO, M. Metadados e recuperação da informação: padrões para bibliotecas digitais. **CIBERÉTICA: SIMPÓSIO INTERNACIONAL DE PROPRIEDADE INTELECTUAL, INFORMAÇÃO e ÉTICA**, v. 2, 2003.

SABBATINI, M. Publicações eletrônicas na internet. São Caetano do Sul: Yendis, 2005. 320 p.

SAMPAIO, P. R. P.; SOUZA, C. A. P. de. Contratos eletrônicos: um novo direito para a sociedade digital? Santa Catarina. **Jornal do Brasil**, 03 fev. 2000. Disponível em: <
<http://www.egov.ufsc.br/portal/sites/default/files/anexos/6070-6062-1-PB.pdf>>. Acesso em: 18 jan. 2013.

SANTAELLA, L. A tecnocultura atual e suas tendências futuras. **Signo y pensamiento**, Bogotá, v.30, n. 60, p. 30-43, enero-junio 2012. Disponível em <
<http://revistas.javeriana.edu.co/index.php/signoypensamiento/article/view/2408/1692>> Acesso em: 12 jan. 2013.

SANTOS, D. S. A. **RDF na interoperabilidade entre domínios na web**. 2002. 108 f. Dissertação (Mestrado em Ciências) – Coordenação de Pós-Graduação em Informática, Universidade Federal de Campina Grande, Paraíba.

SANTOS, P. L. V. A. da C.; ALVES, R. C. V. Metadados e Web Semântica para estruturação da Web 2.0 e Web 3.0. **DataGramZero**, v. 10, n. 6, dez./2009.

Disponível em: < http://www.dgz.org.br/dez09/Art_04.htm>. Acesso em: 12 mar.2013.

SAYÃO, L. F.; MARCONDES, C. H. O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. **TransInformação**, Campinas, v. 20, n. 2, p. 133-148, maio/ago., 2008. Disponível em: < <http://200.20.0.246:8080/jspui/bitstream/1/366/1/Artigo-Marcondes-O%20desafio%20da%20interoperabilidade.pdf> >. Acesso em: 23 fev. 2013.

_____. Bibliotecas digitais e suas utopias. **PontodeAcesso**, Salvador, v. 2, n. 2, p. 2-36, ago./set. 2008. Disponível em: <<http://www.portalseer.ufba.br/index.php/revistaici/article/view/2661/2166>>. Acesso em: 14 fev. 2013.

_____. Afinal, o que é biblioteca digital? **Revista USP**, 2009, n. 80, p. 6-17. Disponível em: <<http://eprints.rclis.org/14675/1/biblioteca-digital.pdf>>. Acesso em: 25 fev. 2013.

SIQUEIRA, M. A. **XML na Ciência da Informação**: uma análise do MARC21. 2003. 133f. Dissertação (Mestrado em Ciência da Informação) – Coordenação de Pós-Graduação em Ciência da Informação, Faculdade de Filosofia e Ciências da Universidade Estadual Paulista, São Paulo.

SILVA, N. C.; SÁ, N. O.; FURTADO, S. R.S. Bibliotecas digitais: do conceito às práticas. In: **Simpósio Internacional de Bibliotecas Digitais**, v. 2, 2005.

SOARES, A. A.A. **Minidicionário da língua portuguesa Soares Amora**. São Paulo: Saraiva, 1997.

SOUZA, R. R.; ALVARENGA, L. Web Semântica e suas contribuições para a ciência da informação. **Ciência da Informação**. , Brasília, v. 33, n. 1, p. 132-141, jan./abril 2004. Disponível em < <http://www.scielo.br/pdf/%0D/ci/v33n1/v33n1a16.pdf> >. Acesso em 15 dez. 2012.

TAMMARO, A. M.; SALARELLI, A. **A Biblioteca Digital**. Brasília: Brique de Lemos, 2008. 378 p.

TENNANT, R. Library catalogs: the wrong solution. **Library journal**, California, feb.2003. Disponível em: <<http://www.libraryjournal.com/article/CA273959.html>>. Acesso em: 16 fev. 2013.

TILLET, B. RDA and the Semantic Web, Linked Data Enviroment. **JLIS.it**, v. 4, n. 1, jan. 2013. Disponível em: <<http://leo.cilea.it/index.php/jlis/article/view/6303/7876>>. Acesso em: 16 fev. 2013.

THOMPSON, H. S. XML Schema Part 1: Structures Second Edition. **W3C Recommendation.(20114) 6-28)[2012-03-28]. hRp: llwww. w3.org/TR/xmlschema- 1**, 2004.

UNICODE, Inc. O que é Unicode?. 2008. Disponível em:<
<http://www.unicode.org/standard/translations/portuguese.html> >. Acesso em: 01 mar. 2013.

YEE, M. M. New perspectives on the shared cataloging environment and a MARC 21 shopping list. Library Resources & Technical Services, v. 48, n. 3, p. 165-178. 2004. Disponível em: <<http://repositories.cdlib.org/postprints/365>>. Acesso em 22 fev. 2013.